23rd ABCM International Congress of Mechanical Engineering
December 6-11, 2015, Rio de Janeiro, RJ, Brazil

PERFORMANCE EVALUATION OF CORRELATIONS TO ESTIMATE THE DIFFUSE COMPONENT FROM GLOBAL IRRADIANCE USING EXPERIMENTAL DATA FROM SONDA STATIONS IN BRAZIL

Alessandro Miranda de Freitas*
alessandroamdf@gmail.com

Willian Mendes Ferreira*
willianm07@gmail.com

Rodrigo Figueiredo Abdo*
rodrigo.fabdo@gmail.com

Matheus Pereira Porto*
matheuspporto@gmail.com

***** Grupos de Energias Renováveis da Escola de Engenharia da UFMG. Laboratório de Termometria – LabTerm. Departamento de Engenharia Mecânica (DEMEC). Universidade Federal de Minas Gerais (UFMG). 6627 Antônio Carlos ave, Pampulha. Belo Horizonte – MG.

This work presents a discussion on the accuracy of the literature correlations to estimate diffuse irradiance from global irradiance, using experimental data set from SONDA stations in Brazil. Monitoring stations to measure direct and diffuse irradiance separately are costly, and for this reason many researchers have used correlations to estimate them indirectly from global irradiance measurements. Despite the fact that these correlations were conceived in most part for Europe and North America, authors have been keeping using them for other locations, as South America for example. Results have shown that literature correlations can estimate with good accuracy the diffuse fraction, for the data set used in this work. Additionally, we suggest more accurate correlations, based on local data set.

Keywords: *solar energy, solar irradiation, fraction diffuse, global irradiance, correlations*

1. INTRODUCTION

The correct evaluation of diffuse and direct components of global radiation is critical for anticipating the thermal performance of various systems, as concentrating solar collectors, in agricultural processes and buildings thermal comfort (Kalogirou, 2009; Duffie and Beckman, 2006). On the other hand, there are several difficulties to estimate the components of solar irradiation by only employing the radiative transfer equation (Muneer et al, 2007), and the estimative become even more difficult when the anisotropic behavior of the atmosphere cannot be accurately provided (Paint et al, 2009). These atmospheric difficulties are mainly due to the clouds, humidity, particles and aerosols, and they all need special and expensive equipment for measurement (Tapakis and Charalambides, 2013).

Devices for global irradiance measurement are more commonly found (because is a cheaper equipment), a well-known alternative is to estimate the diffuse component from the global irradiance. The first empirical correlation that provided a relation between diffuse component and global irradiance was proposed by Liu and Jordan (1960) and has been the most used. Despite the fact it was released for daily basis, it is also used for hourly basis, as function of the clearness index (k_t). Orgill and Hollands (1977) have used data from Canadian stations to estimate hourly fraction diffuse (K_d) in terms of a first polynomial degree. Erbs et al. (1982) introduced a fourth polynomial degree for data from four U.S stations and one Australia's. Reindl et al. (1990) have used an independent data set from the U.S and Europe. Some additional variables were used, such as solar altitude angle (Bugler, 1977; Iqbal et. al., 1983). Bortolini et al. (2012) proposed a multi-location model for forty-four stations in Europe to estimate the K_d.

All existent correlations are based on empirical data, which are associated with specific weather conditions and geographic latitude (Jacovides et al., 2006; Torres et al. 2010). For this particular reason, these correlations have better applicability on local data sets. In this work, we provide a correlation for local monitoring stations in Brazil.

2. SOLAR IRRADIANCE DATA – MEASUREMENTS AND PRE-PROCESSING

2.1 RAW DATA

The data sets used in this work were from five weather stations of the SONDA project (National Organization System of Environmental Data) in Brazil. The data were measured every minute by Kipp & Zonen devices, such as pyranometers, pyrheliometers and shading devices for measuring global radiation, direct and diffuse components.

Table 1 shows the location, period of time, latitude, longitude and altitude of the five weather stations used in this work.

Table 1. Location, period of time, latitude, longitude e altitude of the data sets.

City	Period	Latitude	Longitude	Altitude(m)
Petrolina	Jan 2010-Jan 2015	09° 04' 08" S	40° 19' 11" W	387
Brasília	Jan 2010-Jan 2015	15° 36' 03" S	47° 42' 47" W	1023
Cachoeira Paulista	Mar 2014-Jan 2015	22° 41' 23" S	45° 00' 22" W	574
Curitiba	Sept 2014-Jan 2015	25° 29' 44" S	49° 19' 52" W	891
São Martinho da Serra	Jan 2010-Jan 2015	29° 26' 34" S	53° 49' 23" W	489

The monitoring stations were chosen to represent the latitude range of Brazil, as shows in Fig. 1. Weather stations in the Brazilian northwest do not have devices for measuring direct and diffuse radiation. Hourly direct irradiance on horizontal surface is used to perform quality control.

Figure 1. Distribution of the weather stations considered in this work.

2.2 DIFFUSE FRACTION AND CLEARNESS INDEX

Diffuse radiation is due to the scattering of solar radiation in the atmosphere, and as a consequence the atmosphere composition is critical for the measurements. A form to express atmospheric interference on the radiation is through the hourly diffuse fraction (K_d):

$$K_d = \frac{I_d}{I} \quad (1).$$

Where I_d e I are the hourly diffuse fraction and the hourly global irradiance on a horizontal surface, respectively. The relation between clouds, air mass and the aerosol in the atmosphere with the measured hourly global irradiance can be determined by hourly clearness index (k_t), given by:

$$k_t = \frac{I}{I_0} \quad (2).$$

Where I is the hourly global irradiance and I_0 is the hourly extraterrestrial radiation, which may be determined by:

$$I_0 = \frac{12\times3600}{\pi} G_{sc}\left(1+0.033\cos\frac{360n}{365}\right)\times\left[\cos\varnothing\cos\delta(\sin\omega_2-\sin\omega_1)+\frac{\pi(\omega_2-\omega_1)}{180}\sin\varnothing\sin\delta\right] \quad (3).$$

Where G_{sc} is the solar constant, equivalent to 1367 Wm^{-2} (Iqbal, 1983), $\emptyset$ is local latitude, ω is the angular displacement of the sun according Duffie and Beckman (2006). The subscripts "2" and "1" indicates time sequence. The δ represents the solar declination, i.e. the angular position of the sun at solar noon in relation to the equatorial plane and may be computed by:

$$\delta = 23.45 sin\left(360\frac{284+n}{365}\right) \quad (4).$$

Where n is the day of the year. As the range of K_d and k_t is *[0, 1]*, values that were not in this range were eliminated by quality control algorithms, which is described in the following section. High values of K_d are associated with the predominance of diffuse radiation, typical of overcasting days, leading to low values of k_t. On the other hand, it can be observed that low values of K_d imply clear sky conditions, i.e. high values of k_t. Intermediate values of both may be associated to thin or partially cloudy. As Eq.3 is latitude dependent, Soler (1990) concluded that correlations for K_d and k_t have lower accuracy for wide latitude ranges, as encountered in Brazil.

2.3 QUALITY CONTROL

The correlations accuracy is strongly affected by the quality of the collected data: instabilities, cosine response, misalignment, presence of smoke or dust on the sensors, may mislead the acquisition system to inaccurate data (Geiger et al, 2002; Younes et al., 2005; Muneer and Munawwar, 2006; Journée and Bertrand, 2011).

In this work were used three data control methods simultaneously. The first was based on the proposition of Reindl et al. (1990) that considers conditions totally overcast, at which necessarily $K_d > 0.9$ for $k_t < 0.2$ and totally clear sky conditions where $K_d < 0.8$ for $k_t > 0.6$. The second step is from Europeans Commissions Daylight I - 1993, at which data are rejected when: $0 > I_d > 1.1I$, $0 > I > 1.2I_0$, $0 > I_d > 0.8I_0$, $0 > I_b > I_0$ and $I < 5 W/m^2$. The variable I_b is the hourly beam irradiance on horizontal surface. The third data control is a statistical procedure developed by Muneer et al. (2006). This is an analysis of mean and standard deviation of K_d - $(\overline{K_d}, \sigma_{K_d})$ - concerning the exact division in ten tracks of k_t range. From this, the higher threshold data $(\overline{K_d} + 2\sigma_{K_d})$ and lower threshold data $(\overline{K_d} - 2\sigma_{K_d})$ of each data track were computed and two curves were fitted to form a data envelope, where the data outside were eliminated.

Figure 2 presents in the left column the data set after the first two controls listed above, and the right column after the latter control. Experimental data points on graph K_d - k_t of Fig.2 are similar to the scattered distribution observed by Jacovides et al (1996). Petrolina has low cloudiness, with high concentration in the clear sky region within $0.65 < k_t < 0.75$ (Fig. 2c and Fig.2d). On the other hand, São Martinho da Serra is the opposite, with great part of data close to K_d equals to 1 (Fig.2a and Fig.2b). This fact is explained by the overcasting conditions.

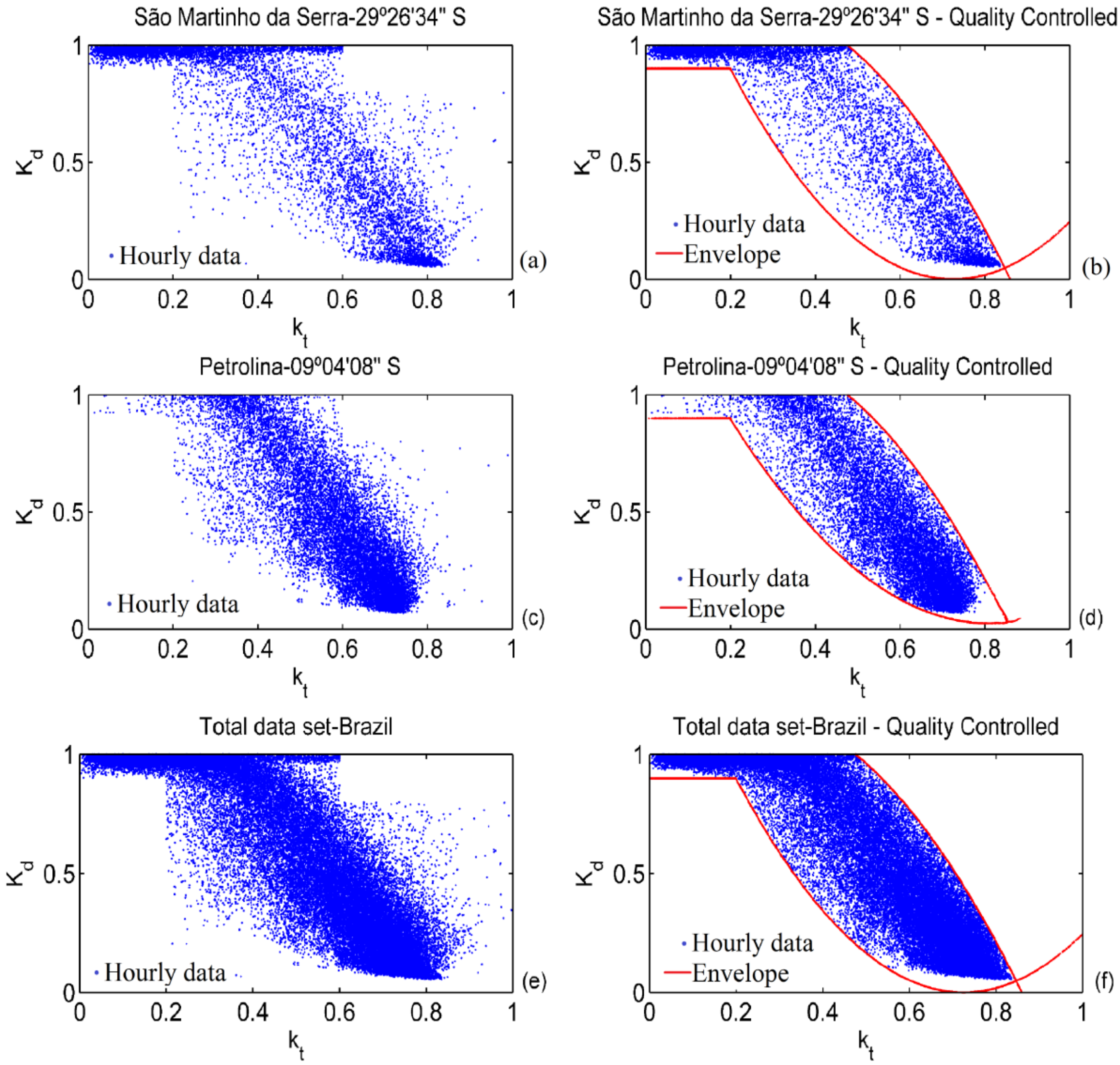

Figure 2. Quality control of data for São Martinho da Serra, Petrolina and total data set (data of five weather stations).

3. RESULTS AND DISCUSSIONS

3.1 LITERATURE CORRELATIONS FOR BRAZIL DATA SET

The classical correlations in the literature were proposed by Orgill and Hollands (1977), Erbs et al (1982) and Reindl et al (1990). Also, we used the proposed correlation by Oliveira et al (2002), which was developed for São Paulo.

In general, K_d is a function of three equations. The first, for the range of k_t < 0.2, and the data can be fitted by a first order polynomial degree. The second, for the range of $0.2 < k_t < 0.75$, which can be adjusted by a first to fourth order polynomial degree. For the latter division, $k_t \geq 0.75$, corresponding to clear sky conditions (with low cloudiness and diffuse concentration), Orgill and Hollands (1977) recommended a constant K_d value for this range, based on observed values from histograms.

For the data set from the weather stations in Brazil, we proposed the following relations:

$$K_d = \begin{cases} -0.1919k_t+0.998 & if \quad k_t < 0.2 \\ 0.9807 - 0.2531\, k_t +2.938\, k_t\text{^}2 -12.26\, k_t\text{^}3 + 9.128k_t\text{^}4 & if \; 0.2 \leq k_t < 0.75 \\ 0.1595 & if \quad k_t \geq 0.75 \end{cases} \tag{5}$$

Figure 3 shows the results for the proposed Eq. (5) and the literature correlations. The presented correlations have similar trends, also close inflection points. In general, differences become more visible at the region between $0.2 < k_t$

< 0.4. In case of Oliveira et al. (2002), differences are more significant, and this is due to the effect of the local altitude (> 700m) and absorption, caused by the high pollution of São Paulo (Oliveira et al., 2002).

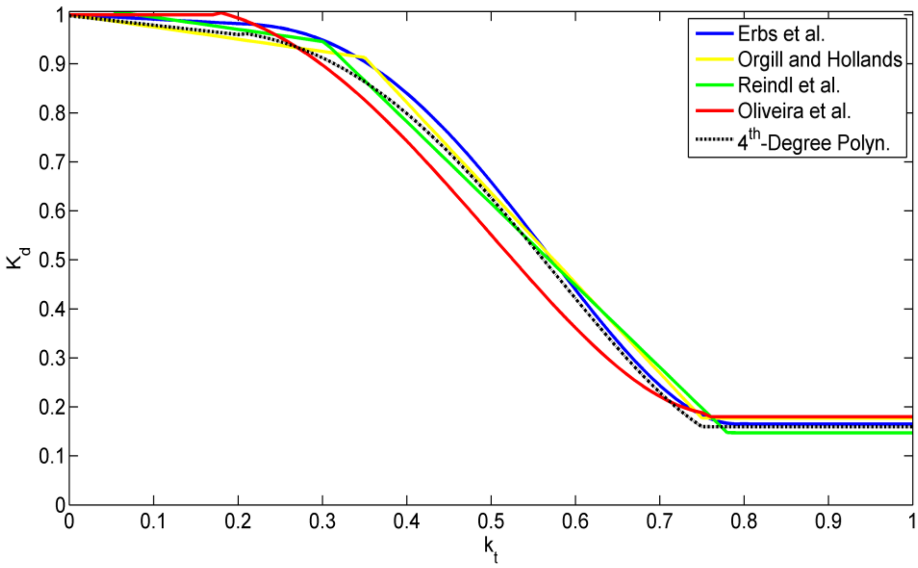

Figure 3. Trends of literature and proposed correlation for the used data set.

3.2 PROPOSED CORRELATIONS FOR EACH WEATHER STATION

In addition to the correlation previously presented (see Eq. 5), we give five other models for each different weather station, as can be seen in Table 2. There are significant differences between the coefficients of each data set, which is depicted in Fig. 4.

Table 2. Proposed correlations for each weather station.

City	Restringe	K_d
Petrolina	$k_t < 0.2$	1 - 0.34k_t
	$0.2 \leq k_t < 0.75$	1.05 - 1.665k_t + 9.61k_t^2 - 24.31k_t^3 + 16.49k_t^4
	$k_t \geq 0.75$	0.1676
Brasília	$k_t < 0.2$	1 – 0.1725k_t
	$0.2 \leq k_t < 0.75$	0.97 + 0.006k_t + 1.13 k_t^2 – 8.801k_t^3 + 7.185k_t^4
	$k_t \geq 0.75$	0.1737
Cachoeira Paulista	$k_t < 0.2$	1 - 0.11k_t
	$0.2 \leq k_t < 0.75$	0.95 + 0.509k_t – 1.309k_t^2 – 2.674k_t^3 + 2.263k_t^4
	$k_t \geq 0.75$	0.1837
Curitiba	$k_t < 0.2$	1 - 0.2205k_t
	$0.2 \leq k_t < 0.75$	0.95 + 0.043k_t + 1.289k_t^2 – 7.851k_t^3 + 5.718k_t^4
	$k_t \geq 0.75$	0.2036
São Martinho da Serra	$k_t < 0.2$	1 - 0.145k_t
	$0.2 \leq k_t < 0.75$	0.99 – 0.528k_t + 4.502k_t^2 – 14.27k_t^3 + 9.662k_t^4
	$k_t \geq 0.75$	0.1662
Total Data Set	$k_t < 0.2$	1 - 0.1919k_t
	$0.2 \leq k_t < 0.75$	0.98 – 0.025k_t + 2.938k_t^2 – 12.26k_t^3 + 9.128k_t^4
	$k_t \geq 0.75$	0.1595

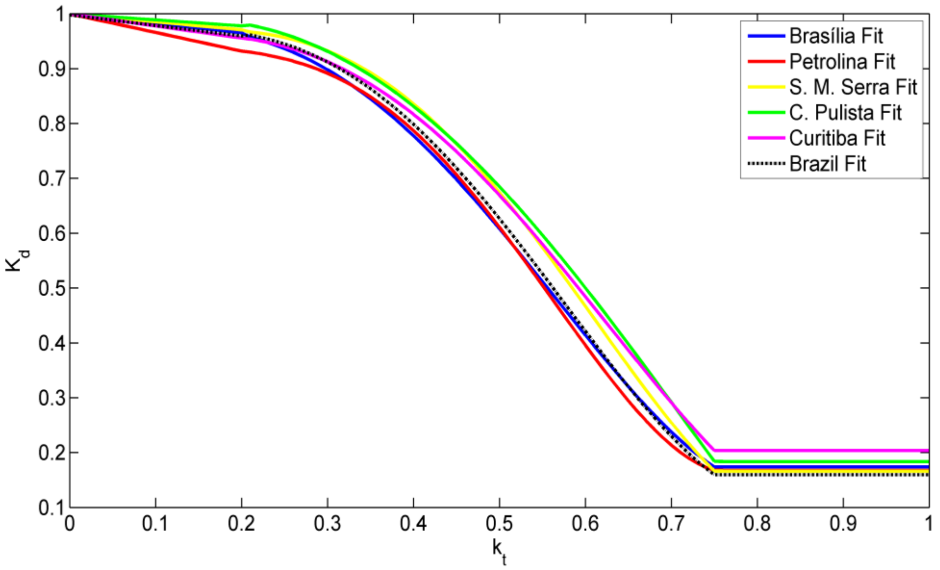

Figure 4. Correlations curves for different cities of Brazil

Similarly to other correlations developed for high altitude (Oliveira et al, 2002; LeBaron and Dirmhim, 1983), it was observed for Brasília (1023 meters) a steeper slope. On the other hand, for Curitiba, located at 891 meters, it was not noticed the same effect. Other effects that may explain this dissimilarity are atmospheric conditions as the turbidity, humidity and aerosol concentration (Gueymard, 1989).

São Martinho da Serra, Cachoeira Paulista and Curitiba have similar trends, and this is due to the small difference in the latitude range (only seven degrees), which implies in closely related air masses and fronts.

For $k_t < 0.2$, Petrolina showed a steeper trend compared to the other cities, also lower values of K_d for the range between $0.2 < k_t < 0.3$. This fact is mainly due to the air masses on the Brazil Northeast, which can be separated into two phenomena. First, the air masses arrive with no humidity, due to physical barriers, creating mostly clear sky atmospheric conditions. Second, in this region occurs a convergence of air masses, leading to the formation of a high-pressure region, at which constrains the clouds formation. Figure 5 is comprised of histograms representing the distribution of points for Petrolina (left side). For this reason, the polynomial fitting to estimate K_d through k_t for Petrolina is difficult, as will be shown in the next section.

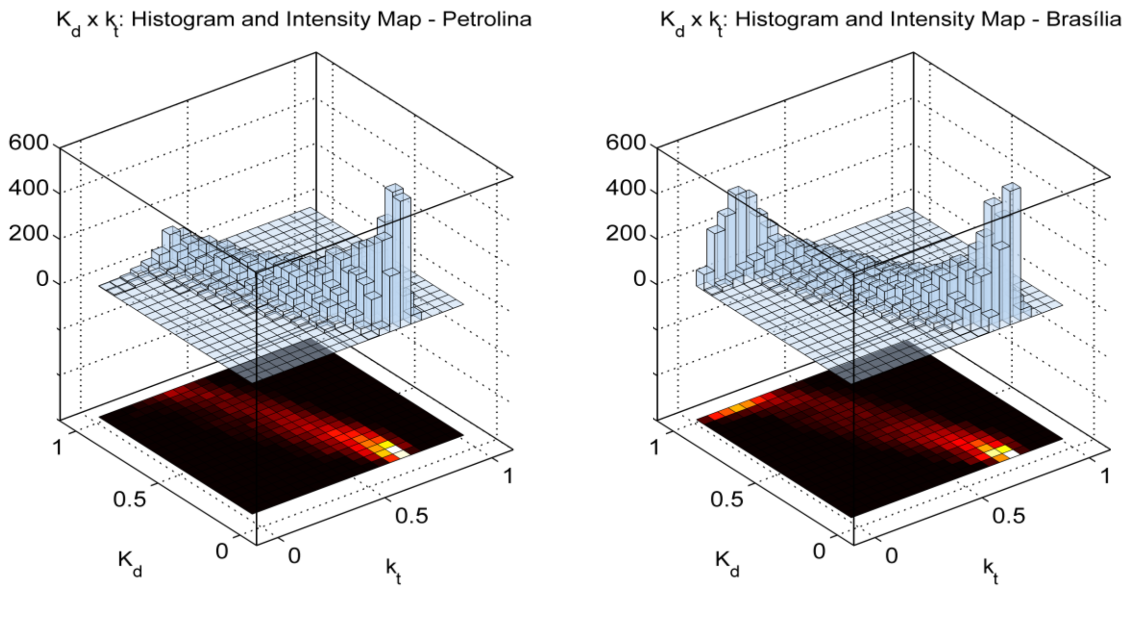

Figure 5. Histogram and map of intensity distribution points of Petrolina and Brasilia.

4. PERFORMANCE EVALUATION

In order to evaluate the applicability of the most classical correlations of literature and the correlations proposed in this paper, we provide *R²* and *RMSE* for each location. Table 3 shows the results for the validity range $0.2 < k_t < 0.75$, for each data set.

Table 3. Statistical evaluation of literature correlations and proposals in this work.

	Orgill and Hollands		**Erbs et al.**		**Reindl et al.**		**Oliveira et al.**		**Proposed correlation**	
DATA SET **(Proposed correlation)**	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*
Petrolina (see Tab. 4)	0.68	0.15	0.71	0.14	0.69	0.15	0.72	0.14	0.73	0.12
Brasília (see Tab. 4)	0.78	0.14	0.84	0.12	0.80	0.13	0.83	0.13	0.85	0.12
Cachoeira Paulista (see Tab. 4)	0.86	0.12	0.87	0.11	0.86	0.11	0.81	0.14	0.89	0.10
Curitiba (see Tab. 4)	0.82	0.12	0.89	0.09	0.85	0.11	0.84	0.12	0.90	0.12
São Martinho da Serra (see Tab. 4)	0.82	0.14	0.88	0.12	0.83	0.14	0.85	0.13	0.89	0.09
Overall data set (see Eq. 5)	0.77	0.14	0.83	0.12	0.78	0.14	0.82	0.13	0.84	0.11

Considering that quality control is performed in accordance with the distribution of each data set, the number of points (k_t, K_d) is different. As a consequence, it is not suggested to compare the performance between the different locations.

As it can be seen on Fig. 2c and Fig. 2d, there is a greater concentration of clear sky conditions for the city of Petrolina, at which is contained in range of $0 < K_d < 0.6$. Therefore, the correlations applied to the data set of this city are associated with low values of R², as shown in Tab. 2. In contrast, São Martinho da Serra atmosphere is highly influenced by air masses, providing changes in humidity, condensation, and cloud formation. As long as there are a great quantity of data in the range of $0.8 < K_d < 1$, a better overall performance was noticed (lower quantity of data is present for lower values of (K_d), as can be seen in the Fig. 2.

Among literature correlations, the fourth polynomial degree of Erbs et al. (1982) presented better statistical performance. However, the correlations proposed in this study improved the statistical performance of the K_d - k_t relation, for the considered data set.

An interesting fact was the poor performance of Oliveira et al model. (2002) for Cachoeira Paulista (R^2 = 0.81). It was due to the distance between these cities (450 *km*). However, it has a R^2 close to the proposed correlation for the overall data set.

5. CONCLUSIONS

This work evaluates classical literature correlations to estimate K_d in term of k_t, also we propose correlations for data sets in Brazil, using weather data available from monitoring stations SONDA. The performance analysis shows that Erbs et al. (1982) obtained the satisfactory results, but the proposed correlation reached an even better performance. We recommend the use of empirical correlations K_d - k_t developed by the nearest station. For this purpose, we also present five correlations addressed for each considered city and one to the entire data set.

It was also concluded that K_d is reasonably correlated to k_t, due to atmospheric instability of the Brazilian tropical climate. Air masses, which are associated to weather variability, are very important to the frequency and presence of the clouds, which has influence on the correlation accuracy. For example, Bortolini et al. (2013), processed data from forty-four stations in Europe, which has a more stable climate, presenting a R^2 of 0.91, while in this work it returned 0.84.

Differences and similarities between all correlations for local data prove the dependency on latitude, which is used as an entry data in Eq. 3. For future works, other parameters such as altitude, pollution and air masses may be incorporated. Another improvement is considering the effects of the cloud type, as highlighted by Oliveira et al. (2002). Particularly, this is an important issue in tropical regions (Webster et al., 1996), where the cloud activity is a dominant characteristic of the local climate.

6. REFERENCES

Bortolini, M., Gamberi, M., Granziani, A., Manzini, R., Mora, C., 2013. "Multi-location model for the estimation of the horizontal daily diffuse fraction of solar radiation in Europe". Solar Energy 67, 208-216.

Bulger, J.W., 1977. "The determination of hourly insolation on an inclined plane using a diffuse irradiance model based on hourly measured global horizontal insolation". Solar Energy 19, 477-491.

Duffie, J.H., Beckman, W.H. "Solar engineering of thermal processe". Wiley Intersciense, New York (2006)

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

Geiger, M., Diabaté, L., Ménard, L., Wald. L., 2002. "A web service for controlling the quality of measurements of global solar irradiation". Solar Energy 73, 475-480.

Gueymard, C.A., 2004. "The sun's total and spectral irradiance for solar energy applications and solar radiation models Solar Energy 76, 423-453.

Iqbal, M. "Introduction to solar radiation". Academic, Toronto (1983).

Jacovides, C.P., Tymvios, F.S., Assimakopoulos, V.D., Kaltsounides, N.A., 2006. "Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. Renew. Energy 31, 2492-2504.

Journée, M., Bertrand, C., 2011. "Quality control of solar radiation data within the RMIB solar measurements network". Solar Energy 85, 72-86.

Kalogirou, S. Solar energy engineering: process and systems/Soteris Kalogirou – 1st ed. (2009)

LeBaron, B., Dirmihim, I., 1983. 'Strengths and limitations of the Liu and Jordan model to determine diffuse from global irradiance". Solar Energy 31, 167-172.

Liu, B.T.U., Jordan, R.C., 1960. "The interrelationship and characteristic distribution of direct, diffuse and total solar radiation". Solar Energy 4, 1-19.

Muneer, T., Munawwar, S., 2006. 'Improved accurancy models for hourly diffuse solar radiation". Journal Solar Energy. T-ASME 128, 104-117

Muneer, T., Younes, S., Munawwar, S., 2007. "Discourses on solar radiation modeling". Renew. Sust Energy Rev. 11, 551-602

Oliveira, A.P., Escobedo, J.F., Machado, A.J., Soares, J., 2002. "Correlation models of diffuse solar radiation applied to the city of São Paulo, Brazil". Appl. Energy 71, 59-73.

Orgill, J.F., Hollands, K.G.T., 1977. "Correlation equation for hourly diffuse radiation on a horizontal surface". Solar Energy 19, 357-359.

Pinte, C., Harries, T.J., Min, M., Watson, A.M., Dulleomond, C.P., Woitke, P., Ménard, F., Duran Rojas, M.C., 2009. "Benchmark problems for continuum radiative transfer, High optical depths anisotropic scattering and polarization". Astrom. Astrophys. 1, 1-15.

Reindl, D.T., Beckman, W.A., Duffie, J.A., 1990. "Diffuse fraction correlation" Solar Energy 45, 1-7.

Soler, A., 1990. "Dependence on latitude of the relation between the diffuse fraction of solar radiation and the ratio of global-to-extraterrestrial radiations for monthly average daily values. Solar Energy 44, 297-302

Tapakis, R., Charalambides, A.G., 2013. "Equipment and methodologies for cloud detection and classification:a review". Solar Energy 95, 392-430

Torres, J.L., de Blas, M., Garcia, A., de Francisco, A., 2010. "Comparative study of various models in estimating hourly diffuse solar irradiance". Renew. Energy 35, 1325-1332.

Webster, P.J., Clayson, C.A., Curry, J.A., 1996. "Clouds, Radiations, and the Diurnal Cycle of Sea Surface Temperature in the Tropical Western Pacific". Journal Climate 9, 1712-1730

Younes, S., Claywell, R., Muneer, T., 2006. "Quality control of solar radiation data:present status and proposed new approaches". Energy 30, 1533-1549.