Theor. Appl. Climatol. 75, 213–224 (2003)
DOI 10.1007/s00704-003-0733-9

Theoretical and Applied Climatology
Printed in Austria

[1] Department of Geography, Campus of Jataí, Federal University of Goiás, Jataí, Goiás, Brazil
[2] Department of Natural Resources, School of Agronomic Science, UNESP, Botucatu, São Paulo, Brazil
[3] Group of Micrometeorology, Department of Atmospheric Sciences, Institute of Astronomy, Geophysics and Atmospheric Sciences, University of São Paulo, São Paulo, Brazil

Modelling frequency distributions of 5 minute-averaged solar radiation indexes using Beta probability functions

H. F. Assunção[1], **J. F. Escobedo**[2], and **A. P. Oliveira**[3]

With 6 Figures

Received September 2, 2002; revised February 28, 2003; accepted March 16, 2003
Published online September 10, 2003

Summary

Five minute-averaged values of sky clearness, direct and diffuse indices, were used to model the frequency distributions of these variables in terms of optical air mass. From more than four years of solar radiation observations it was found that variations in the frequency distributions of the three indices of optical air mass for Botucatu, Brazil, are similar to those in other places, as published in the literature. The proposed models were obtained by linear combination of normalized Beta probability functions, using the observed distributions derived from three years of data. The versatility of these functions allows modelling of all three irradiance indexes to similar levels of accuracy. A comparison with the observed distributions obtained from one year of observations indicate that the models are able to reproduce the observed frequency distributions of all three indices at the 95% confidence level.

List of symbols

Symbol	Description
a, b, c	Polynomial coefficients
k_t	Clearness index (dimensionless)
k_{bh}	Horizontal direct index (dimensionless)
k_d	Diffuse index (dimensionless)
G_0	Extraterrestrial solar irradiance ($\mathrm{W\,m^{-2}}$)
G	Global solar irradiance ($\mathrm{W\,m^{-2}}$)
G_b	Normal beam solar irradiance ($\mathrm{W\,m^{-2}}$)
G_{bh}	Direct solar irradiance on horizontal surface ($\mathrm{W\,m^{-2}}$)
G_d	Diffuse solar irradiance ($\mathrm{W\,m^{-2}}$)
m_a	Relative relative optical air mass (dimensionless)
$\bar{m}_a$	Mean of relative relative optical air mass (dimensionless)
Δz	Altitude (m)
θ_z	Zenith angle of the sun (degrees)
f_i	Observed frequency in interval i
s^2	Variance of frequency distribution
$\bar{x}$	Mean of frequency distribution
x_i	Central value
n	Number of intervals
$\Gamma(\cdot)$	Gamma function
α	Parameter of Beta function
β	Parameter of Beta function
τ	Normalization constant
$f(\cdot)$	Probability function
$F(\cdot)$	Cumulative distribution function
$I(\cdot)$	Incomplete Beta function
$f_1(\cdot)$, $f_2(\cdot)$	Probability function for first and second peaks
$P[\cdot]$	Occurrence probability of generic irradiance indexes
$g(\cdot)$	Polynomial function
$\bar{k}_t$, $\bar{k}_{t_1}$, $\bar{k}_{t_2}$	Experimental means of k_t for full, first peak and second peak of frequency distribution
s_1^2, s_2^2	Experimental variances of irradiance indexes for first peak and second peak of frequency distribution

$\hat{\bar{\mathbf{k}}}_t$, $\hat{\bar{\mathbf{k}}}_{t_1}$, $\hat{\bar{\mathbf{k}}}_{t_2}$	Estimated means of clearness index with polynomial regression for full, first peak and second peak of frequency distribution
$\hat{s}^2$, $\hat{s}_1^2$, $\hat{s}_2^2$, $\hat{s}_3^2$	Estimated irradiance indexes variances with polynomial regression for full, first, second and third peaks of frequency distribution
k_{bt}	Transformed scale for k_{bh}
τ_1, τ_2, τ_3	Normalization constants for k_{bh} frequency distribution
$\bar{\mathbf{k}}_{bt}$, $\bar{\mathbf{k}}_{bt_1}$, $\bar{\mathbf{k}}_{bt_3}$	Experimental means of transformed scale of k_{bh} for full, first peak and third peak of frequency distribution
$\hat{\bar{\mathbf{k}}}_{bt}$, $\hat{\bar{\mathbf{k}}}_{bt_1}$, $\hat{\bar{\mathbf{k}}}_{bt_3}$	Estimated means of transformated scale of k_{bh} with polynomial regression for full, first peak and third peak of frequency distribution
$\bar{k}_d$	Experimental mean of k_d for full frequency distribution
$\hat{\bar{k}}_d$	Estimated mean of k_d with polynomial regression for full frequency distribution
MBE	Mean bias error
RMSE	Root mean square error

1. Introduction

Understanding short term variations in solar radiation has become important for many solar energy applications. Variations course a significant impact on the performance of non-linear and fast response devices such as converter systems of solar energy (Tovar et al., 1998). With the increase in the sensitivity of solar radiation sensors they are now capable of measuring fluctuations on short time scales, mostly due to clouds, assessing solar system performance, photosynthetic production in plants (Berninger, 1994). In addition, estimates of solar irradiance at the surface by remote sensing techniques have become more efficient (Tovar et al., 1998).

The first numerical models developed to assess the performance of solar energy collectors used the concept of solar energy utilizability. This concept was introduced at the end of the 1950s as a tool to estimate the efficiency of solar energy collectors (Jurado et al., 1995). First, the utilizability coefficient was based on hourly values of solar radiation and on daily estimates of the clearness index. Later, this coefficient was obtained from the cumulative distribution function of solar radiation received on a horizontal surface, conditioned by the sunshine time fraction (Suehrcke and McCormick, 1989). Due to lack of sophisticated instruments this later approach was used over a long period of time.

One of the first attempts to take into consideration higher frequency variability of solar radiation was carried out by Liu and Jordan (1960). They determined a set of frequency distributions for daily values of the clearness index, global solar radiation divided by extraterrestrial solar radiation at the top of the atmosphere. These depended only on the monthly-averaged clearness index. These functions were considered as independent of location and time. Later, several other worker determined that these functions were not as universal as initially thought. As a consequence, all frequency distributions have to be considered dependent on parameters such as: latitude, albedo, optical air mass, etc (Jurado et al., 1995).

About 30 years later, Suehrcke and McCormick (1988a), using solar radiation data, sampled over one minute-intervals in Western Australia and found out that the clearness index has a bimodal distribution, instead of the normal distribution assumed in previous studies. The bimodal distribution was appropriately adjusted by a probability function derived from the Boltzmann statistics (Suehrcke and McCormick, 1988a). In this derivation, atmospheric variations were taken into account by the optical air mass and the clearness index.

The variability of clearness, diffuse and direct indices are important parameters not only for the above mentioned practical applications, but also for climate characterization studies. These indices portray the dominant radiometric properties of the atmosphere for solar radiation.

The statistical behavior of short term values of solar radiation in several other places of the world has been studied by Suehrcke and McCormick (1988a); Suehrcke and McCormick (1988b); Skartveit and Olseth (1992); Gansler et al. (1995); Jurado et al. (1995); Tovar et al. (1998); Tovar et al. (1999). These papers all revealed a bimodal distribution for the clearness index. Jurado et al. (1995) found that solar radiation measurements sampled over intervals smaller than 60 minutes, also have a bimodal character.

The main goal of this study is to model the frequency distributions of clearness, direct and diffuse indices for short term variations of solar radiation at the surface, in terms of optical air mass. The frequency distributions of the irradiance indices are estimated using five minute-averaged values observed during four years in the rural area of Botucatu, Brazil.

2. Methodology

The data set consists of global (G) and direct beam (G_b) solar irradiance values at the surface. These were sampled with 0.2 Hz frequency and averaged over five minute-periods by an automatic data acquisition system, CR23X, manufactured by Campbell Scientific Inc. Global solar irradiance was measured with a Spectral Precision Pyranometer manufacture by Eppley Lab Inc. The beam irradiance was measured with a Normal Incidence Pyrheliometer coupled with a ST-3 solar tracker, both also from Eppley. These sensors were regularly calibrated using a reference specially designed for this purpose.

Measurements were carried out continuously from March 1996 to March 2000 at a radiometric station located in the rural area of Botucatu, Brazil (22°51′ S, 48°26′ W, 786 m above m.s.l.). As in the paper by Tovar et al. (1998), the data set was divided into two parts. The first three-fourth was used to develop the model and the last one-fourth for validation. To avoid errors related to cosine effects, only measurements corresponding to solar zenith angles (θ_z) smaller than 75° have been selected.

The horizontal direct solar irradiance (G_{bh}) was computed by projecting the direct beam irradiance on the direction normal to the surface:

$$G_{bh} = G_b \cos\theta_z \tag{1}$$

The diffuse component of solar irradiance (G_d) was computed as the difference of global and beam components.

$$G_d = G - G_{bh} \tag{2}$$

The irradiance indices used here are the clearness index (k_t), the direct index (k_{bh}) and the diffuse index (k_d). They were evaluated according to:

$$k_t = G/G_0 \tag{3}$$

$$k_{bh} = G_{bh}/G_0 \tag{4}$$

Table 1. Grouping of mean values for optical air mass ($\bar{m}_a$) in the range of the observed variation

Class interval	Central value – $\bar{m}_a$
$0.9 \le m_a \le 1.1$	1.0
$1.4 \le m_a \le 1.6$	1.5
$1.8 \le m_a \le 2.2$	2.0
$2.7 \le m_a \le 3.3$	3.0

$$k_d = G_d/G_0 \tag{5}$$

where G_0 is the extraterrestrial solar irradiance calculated according to Iqbal (1983). The clearness index is related to the other two indices by the following relation $k_t = k_{bh} + k_d$.

The relative optical air mass (m_a) was estimated from:

$$m_a = \sec\theta_z \cdot e^{-0.0001184\Delta z} \tag{6}$$

where Δz is the station altitude in meters above mean sea level.

The values of each irradiance index were grouped in classes of relative optical air mass, identified by central values according to Suehrcke and McCormick (1988a) as described in Table 1. Each class was divided into 50 intervals of size 0.02 (Tovar et al., 1998; Tovar et al., 1999).

The frequency distributions of k_t, k_{bh} and k_d according to the relative optical air mass were fitted by the Beta probability function defined as:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{7}$$

Where argument x corresponds to the three irradiance indices.

The parameters α and β are positive and can be computed from the following expressions:

$$\beta = \left(\frac{\bar{x}^3 - 2\bar{x}^2 + \bar{x}}{s^2}\right) + \bar{x} - 1 \tag{8}$$

$$\alpha = \left(\frac{\bar{x}^2 - \bar{x}^3}{s^2}\right) - \bar{x} \tag{9}$$

where

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} \tag{10}$$

$$s^2 = \frac{\sum_{i=1}^{n} f_i x_i^2 - \bar{x}^2 \sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} f_i - 1} \tag{11}$$

are, respectively, the mean and the variance for each frequency distribution.

In (10) and (11), f_i is the observed frequency in the *ith* interval, n is the number of intervals (equal to 50) and x_i is the central value of each class interval. $\Gamma(\gamma)$ is the gamma function given by:

$$\Gamma(\gamma) = \sqrt{\frac{2\pi}{\gamma}}\, e^{\gamma}\left[\ln(\gamma) - \left(1 - \frac{1}{12\gamma^2} + \frac{1}{360\gamma^4} - \frac{1}{1260\gamma^6}\right)\right] \tag{12}$$

The probabilistic behavior of k_t, k_{bh} and k_d, can be characterized by the cumulative distribution function $F(x, t)$ and is associated to the probability function $f(x, t)$. In our case, the dependence on time t is implicitly represented by the relative optical air mass, so that:

$$F(x, m_a) = P[x(m_a) \leq x] \tag{13}$$

where $P[x(m_a) \leq x]$ is the occurrence probability of the irradiance index of a certain relative optical air mass.

Each function obeys the following normalization condition:

$$F(x, m_a) = \int_0^1 f(x, m_a)dx = 1 \tag{14}$$

The Eqs. (13) and (14) are used to define the utilizability coefficient in terms of relative optical air mass (Suehrcke and McCormick, 1989). The relative cumulative frequency represents the fraction of time in which the stochastic variable is located below a given value at each five minute interval.

3. Frequency distribution of the clearness index

Figure 1 shows the relative frequency distribution of the clearness index for the relative optical air mass intervals 1.0, 1.5, 2.0 and 3.0, obtained from the first three-fourth period of the data set. A bimodal distribution can be observed for all intervals of m_a. According to Tovar et al. (1998) the presence of two maxima indicates that each frequency distribution of k_t can be related to two different levels of solar irradiance for each relative optical air mass interval. A flat maximum exists at low levels of k_t, corresponding to cloudy conditions. The second steep peak at high levels of k_t, is related to clear sky conditions.

The amplitude of the second peak decreases as the optical air mass increases so that at higher optical air masses the bimodal characteristic of the distribution disappear. Similar behavior was found by Skartveit and Olseth (1992) and Gansler et al. (1995). On the other hand,

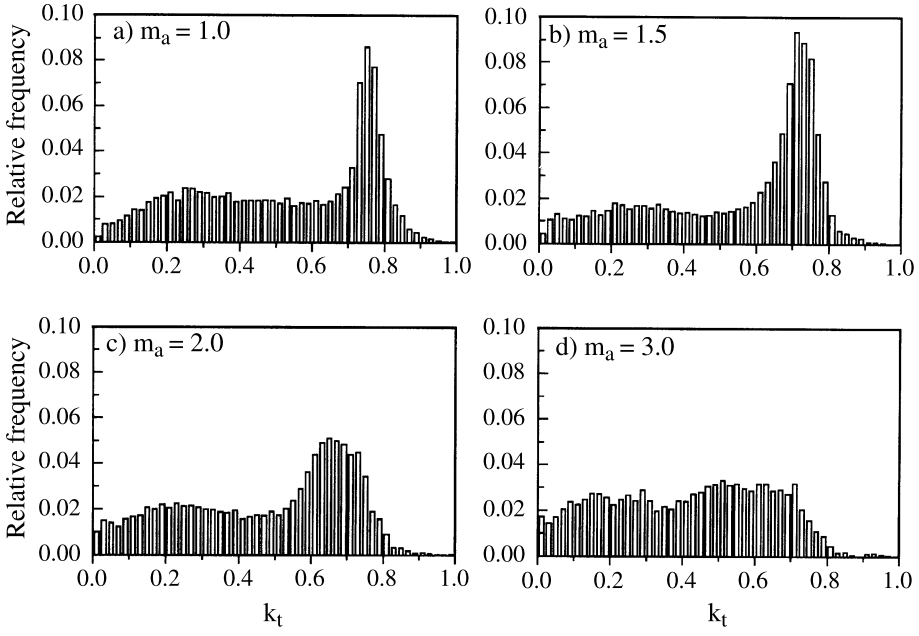

Fig. 1. Relative frequency distribution of k_t for **a**) $m_a = 1.0$, **b**) $m_a = 1.5$, **c**) $m_a = 2$ and **d**) $m_a = 3.0$. Observations carried out in the first period (March 1996 to 1999)

Table 2. Geographic position and data characteristics of the sites used to compare the frequency distribution of clearness, direct and diffuse index

Site	Geographic position	Altitude (m above m.s.l.)	Period of observation (year)	Interval of sampling (minutes)
Botucatu, Brazil	22°51′ S, 48°26′ W	786	4	5
Atlanta, USA	33.65° N, 84.43° W	315	1	
San Antonio, USA	29.46° N, 98.49° W	242		5
Geneva, USA (Skartveit and Olseth, 1992)	46.20° N, 06.15° E	370		
Atlanta, USA	33.65° N, 84.43° W	315	1	
San Antonio, USA	29.46° N, 98.49° W	242		1
Albany, USA (Gansler et al., 1995)	42.70° N, 42.75° W	17		
Perth, Australia (Suerhcke and McCormick, 1988a)	31.93° S, 115.83° W	17	1	1
Seville, Spain (Jurado et al., 1995)	37.40° N, 6.08° W	12	8	5
Granada, Spain (Tovar et al., 1998)	37.13° N, 3.63° E	687	3	1

Suerhcke and McCormick (1988a), Jurado et al. (1995) and Tovar et al. (1998) found that the bimodal character of the k_t frequency distribution becomes more distinct as the optic air mass increases. The geographic position, data sample interval, and information about the above mentioned studies are given in Table 2. Even though the pattern found for the k_t frequency distribution in Botucatu is similar to the one observed in USA and Switzerland studies, there is no apparent relation between these places (Table 2). This behavior cannot be attributed to similarity of climate, where k_t frequency distribution matches. For instance, Botucatu climate is classified according to the methodology of Köppen, as Cwa, a wet and cloudy summer, between December and March, and a dry and cloudless winter, between June and September. The annual mean temperature (relative humidity) is 20 °C (55%), ranging from a maximum in the summer of 22.5 °C (80%), to a minimum in the winter of 17.5 °C (30%). The other areas shown in Table 2 have a totally different climate.

Although several factors can influence the exact form of the of k_t distribution, these effects can be associated with the crescent scattering of direct irradiance due to the increase in atmospheric optical depth. They may also be associated with an increase in the shadow area cast by clouds on the surface, and to an effective thickening of the cloud layers. All these effects reduce the transparency of the atmosphere, thereby favoring the increase of diffuse solar radiation at the surface. Suehrcke and McCormick (1988b) suggested that non zero frequency for $k_t > 0.905\mathrm{e}^{-0.074 m_a}$ is caused by spurious intensification of solar radiation at the surface and by multi-reflection between tower clouds (Fig. 1a and 1b).

In the model proposed here, the frequency distribution is reproduced by the linear combination of two Beta probability functions:

$$f(k_t, m_a) = \tau \cdot f_1(k_t) + (1 - \tau) \cdot f_2(k_t) \qquad (15)$$

where $f_1(k_t)$ and $f_2(k_t)$ are probability functions as given by Eq. (7), and τ is the normalization constant, ranging between 0 and 1, and estimated as:

$$\tau = \frac{(\bar{k}_{t_2} - \bar{k}_t)}{(\bar{k}_{t_2} - \bar{k}_{t_1})} \qquad (16)$$

where $\bar{k}_t$, $\bar{k}_{t_1}$ and $\bar{k}_{t_2}$ are, respectively, the means in the whole frequency distribution domain, in the first peak and in the second peak.

The Beta cumulative distribution function cannot be expressed analytically. However, Eq. (14)

Table 3. Frequency distribution parameters of k_t for different values of relative optical air mass

$\bar{m}_a$	$\bar{k}_t$	$\bar{k}_{t_1}$	$\bar{k}_{t_2}$	s_1^2	s_2^2
Botucatu – Brazil					
1.0	0.5640	0.3593	0.7589	0.0301	0.0024
1.5	0.5206	0.3053	0.7234	0.0272	0.0026
2.0	0.4650	0.2371	0.6650	0.0162	0.0068
3.0	0.4147	0.1709	0.5806	0.0085	0.0136
Granada – Spain (Tovar et al., 1998)					
1.0	0.684	0.34	0.75	–	–
1.5	0.637	0.31	0.73	–	–
2.0	0.598	0.31	0.71	–	–
3.0	0.526	0.25	0.63	–	–

Table 4. Coefficients to estimate means and variances of k_t frequency distributions, in function of relative optical air mass, using Eq. (17)

$g(m_a)$	a	b	c	R^2
Botucatu – Brazil				
$\hat{\bar{k}}_t$	0.0292	– 0.1995	0.7499	0.998
$\hat{\bar{k}}_{t_1}$	0.0341	– 0.2403	0.5844	0.998
$\hat{\bar{k}}_{t_2}$	0.0078	– 0.1272	0.8916	0.998
$\hat{s}_1^2$	0.0027	– 0.0027	0.0535	0.966
$\hat{s}_2^2$	0.0015	0.0015	0.0001	0.983
Granada – Spain (Tovar et al., 1998)				
$\hat{\bar{k}}_t$	–	–	–	–
$\hat{\bar{k}}_{t_1}$	0.010	– 0.0954	0.4692	0.992
$\hat{\bar{k}}_{t_2}$	– 0.012	– 0.0152	0.7628	0.996
$\hat{s}_1^2$	–	–	–	–
$\hat{s}_2^2$	–	–	–	–

can be solved numerically using Incomplete Beta functions or using a Beta cumulative distribution function, available in most software (MS Excel, Mathcad, etc.).

It is interesting to note that the variation of model parameters with relative optical air mass are able to reproduce the major features of the observed relative frequency (Fig. 1). For instance, the mean values of k_t are displaced towards low values as the relative optical air mass increases (see Table 3). According to Skartveit and Olseth (1992), this reflects the fact that an increase in diffuse irradiance partly compensates an increase in the extinction by aerosols and thin clouds, leaving the global irradiance significantly less affected than the beam irradiance.

In fact, the parameters in Table 3 can be used to describe the probability function of the clearness index in terms of the relative optical air mass. A continuous expression relating these parameters to m_a would give a more complete physical description of this distribution. This procedure would improve the model and allow its application to a wider range of relative optical air masses.

The expressions relating the model parameters to m_a were obtained by adjusting a 2nd order polynomal to the experimental frequency distribution, observed in the first three-fourth period, by a linear regression method. The best fit for $\bar{k}_t$, $\bar{k}_{t_1}$, $\bar{k}_{t_2}$, s_1^2 and s_2^2 was obtained by the following polynomial:

$$g(m_a) = a \cdot m_a^2 + b \cdot m_a + c \qquad (17)$$

where a, b and c are the coefficients of $\hat{\bar{k}}_t$, $\hat{\bar{k}}_{t_1}$, $\hat{\bar{k}}_{t_2}$, $\hat{s}_1^2$ and $\hat{s}_2^2$ obtained from Eq. (17) as given on Table 4.

The parameters $\bar{k}_{t_1}$ and $\bar{k}_{t_2}$ have similar trends to those obtained by Tovar et al. (1998), using solar irradiances measured every single minute in Granada, Spain. However, these differed in magnitude (Table 3), possibly due to climate or data sampling differences (Table 2).

In order to validate the model, the analytical probability functions (7) built up using Eq. (17) were compared to the experimental frequency distribution in the second one-fourth period of observations (Fig. 2). Figure 2 displays the analytical probability function of k_t for the observations carried out in the second period (March 1999 to 2000), for $m_a = 1.0$ and $m_a = 2.0$. As it will be discussed later, the model derived here reproduced the observations quite well.

This result agrees with the one obtained by Suehrcke and McCormick (1988a); Skartveit and Olseth (1992); Jurado et al. (1995); and Tovar et al. (1998). These authors introduced a model to reproduce analytically the observed bimodal character of these distributions. For instance, Suehrcke and McCormick (1988a) proposed a model using three functions of three levels of solar radiation. Two functions represent the distribution k_t under totally clear and totally cloudy conditions, and the third one reproduces the distribution of k_t in an intermediate region.

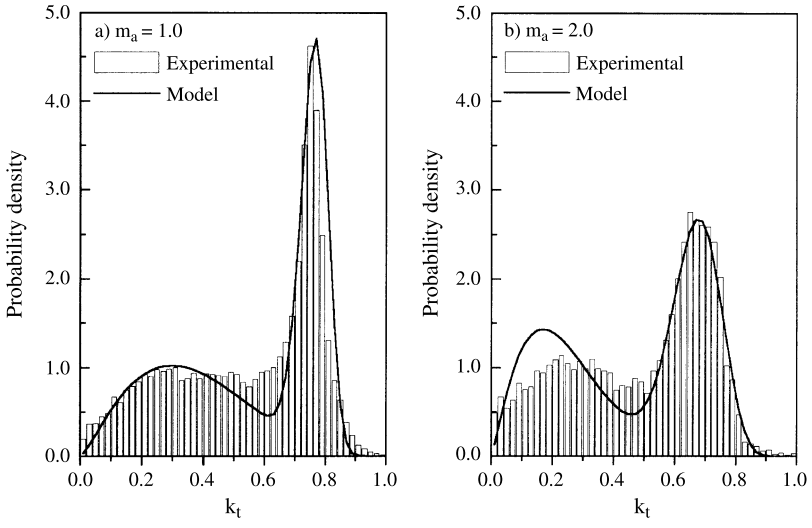

Fig. 2. Probability functions of k_t modeled for **a**) $m_a = 1.0$ and **b**) $m_a = 2$. Observations were carried out in the second period (March 1999 to 2000)

Skartveit and Olseth (1992) used a linear combination of two Beta probability functions; Jurado et al. (1995) combined two Gauss functions; while Tovar et al. (1998) used two logistic probability functions derived from Boltzmann statistics.

4. The frequency distribution of the horizontal direct index

Figure 3 presents the frequency distribution of the direct index for different relative optical masses, showing the expected two peak shape (Skartveit and Olseth, 1992; Tovar et al., 1999). In order to develop a model for this distribution it was assumed that it is a linear combination of three distinct Beta probability functions: exponential, uniform and asymmetric.

The exponential range ($0.00 \leq k_{bh} < 0.05$) corresponds to atmospheric conditions in which direct solar irradiance is close to zero. These conditions range from a sky totally covered by thick layers of clouds to one that is partially covered, where transient clouds hide the sun for short periods. Different to the results obtained by Skartveit and Olseth (1992) and by Tovar et al.

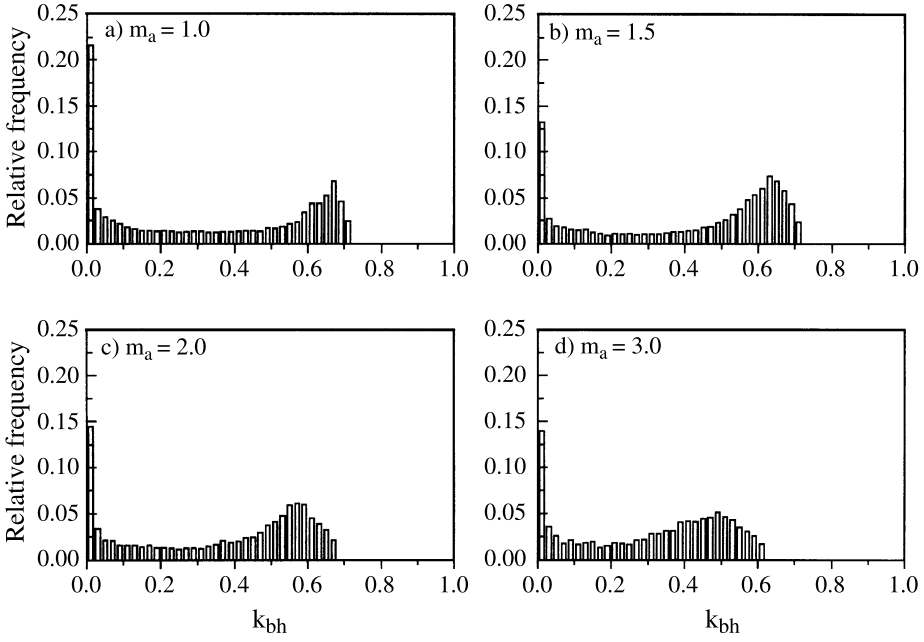

Fig. 3. Relative frequency distribution of k_{bh} for **a**) $m_a = 1.0$, **b**) $m_a = 1.5$, **c**) $m_a = 2$ and **d**) $m_a = 3.0$. Observations carried out in the first period (March 1996 to 1999)

(1999), the frequencies found here are more pronounced at small values of relative optical mass (Fig. 3a). The intensification of *cumulus* convection around noon-time, which is more often the case in tropical regions, may explain this discrepancy.

The uniform range is characterized by an approximately constant level of relatively small frequencies of k_{bh}. These results indicate that the frequency distribution of k_{bh} do not vary considerably with respect to the relative optical air mass. They also match with the frequency distribution of k_{bh} found by Skartveit and Olseth (1992) and Tovar et al. (1999). The behavior of k_{bh} in this fashion is due to the atmospheric conditions when the sky is covered by thin layers of clouds such as *cirrus* and *cirrustratus*.

The asymmetric range is characterized by a large region with progressive and smooth growth in the relative frequency of k_{bh}, and by a peak followed by a region with a sharp decrease in the relative frequency of k_{bh}. The variation in this later region decreases as the relative optical mass increases. In this case, the atmospheric condition is characterized by a completely clear sky. Under these circumstances, direct solar radiation is mainly attenuated by scattering (Skartveit and Olseth, 1992).

In a relative frequency distribution, k_{bh} value shifts from 0.75 ($m_a = 1.0$, Fig. 3a) to 0.61 ($m_a = 3.0$, Fig. 3d). According to Skartveit and Olseth (1992), this shift is caused by attenuation of the direct solar beam due to ozone, water vapor and atmospheric turbidity.

The analysis described above confirms that the probability density distribution of the direct index can be reproduced by combining exponential, uniform and asymmetric Beta probability functions. The shape of each distribution is given by three different sets of the α and β coefficients. For $\alpha < 1$ and $\beta \geq 1$, the distribution has an exponential shape. For $\alpha = 1$ and $\beta = 1$, the distribution is uniform, and for $\alpha > \beta$ the distribution is asymmetric. Skartveit and Olseth (1992) used similar Beta probability functions to fit the observations, while Tovar et al. (1999) used a logistic modified function. In this later work, frequencies of k_{bh} in the interval (0.00–0.02) have been selected, and the distribution was fitted by an exponential function of the relative optical mass. In this case, a second interval ($0.02 < k_{bh} < 1.00$) was considered to fit a modified logistic function.

In order to accommodate the observed behavior of the relative frequency of k_{bh}, a scale transformation for k_{bh} has been introduced. The transformed k_{bh} is indicated by k_{bt} and calculated as:

$$k_{bt} = \frac{k_{bh}}{0.82 - 0.07 m_a} \tag{18}$$

where the denominator of Eq. (18) represents the maximum expected value of k_{bh} for relative optical air mass m_a.

The resulting $f(k_{bt}, m_a)$, obtained by linear combination of the three Beta probability functions $f(k_{bt})$ are given by:

$$f(k_{bt}, m_a) = \sum_{i=1}^{3} \tau_i f_i(k_{bt}) \tag{19}$$

The cumulative distribution function $F(k_{bt}, m_a)$ in this case becomes:

$$F(k_{bt}, m_a) = \sum_{i=1}^{3} \tau_i I(\alpha_i, \beta_i, k_{bt}) \tag{20}$$

where $I(\alpha_i, \beta_i, k_{bt})$ is an incomplete Beta function for each distribution of k_{bt}. In the second distribution function, $\alpha_2 = 1$ and $\beta_2 = 1$ was assumed. In this, τ_i are weight parameters related to each Beta probability function i. They are indicated by the following expressions:

$$\tau_1 = \frac{(\bar{k}_{bt_3} - \bar{k}_{bt})}{2(\bar{k}_{bt_3} - \bar{k}_{bt_1})} \tag{21}$$

$$\tau_2 = 1 - \bar{k}_{bt} \tag{22}$$

$$\tau_3 = (1 - \tau_1 - \tau_2) \tag{23}$$

where $\bar{k}_{bt}$ is a mean value of k_{bt} in the entire distribution domain; $\bar{k}_{bt_1}$ and $\bar{k}_{bt_3}$ are, respectively, the mean values of the exponential and asymmetric distribution.

Table 5 shows the frequency distribution parameters for k_{bt}, in the various relative optical air mass intervals based on observations carried out in the first three-fourth period. The 2nd order polynomials that best fit the experimental distribution parameters from Table 5, are:

$$g(m_a) = a \cdot m_a^2 + b \cdot m_a + c \tag{24}$$

Table 5. Frequency distribution parameters of k_{bh} for different values of relative optical air mass

m_a	$\bar{k}_{bt}$	$\bar{k}_{bt_1}$	s_1^2	$\bar{k}_{bt_3}$	s_3^2
1.0	0.475	0.0236	0.00034	0.8765	0.0072
1.5	0.579	0.0245	0.00036	0.8470	0.0091
2.0	0.539	0.0264	0.00040	0.7954	0.0156
3.0	0.507	0.0306	0.00054	0.7298	0.0228

Table 6. Coefficients to estimate means and variances of k_{bh} frequency distributions, in function of relative optical air mass, using Eq. (24)

$g(m_a)$	a	b	c	R^2
$\hat{\bar{k}}_{bt}$	−0.07510	0.31300	0.23880	0.96780
$\hat{\bar{k}}_{bt_1}$	0.00070	0.00080	0.02180	0.99875
$\hat{\bar{k}}_{bt_3}$	0.00990	−0.11950	0.99840	0.99745
$\hat{s}_1^2$	0.00004	−0.00005	0.00030	0.99865
$\hat{s}_3^2$	−0.00050	0.01050	−0.00440	0.99710

In this case, $g(m_a)$ is a function to calculate approximately $\hat{\bar{k}}_{bt}$, $\hat{\bar{k}}_{bt_1}$, $\hat{\bar{k}}_{bt_3}$, $\hat{s}_1^2$ and $\hat{s}_3^2$ values, with which α and β were estimated. The coefficients a, b and c are given in Table 6. The accuracy of these expressions is equivalent to the polynomial described for the clearness index in the previous section.

Figure 4 displays comparisons between model predictions and the experimental frequency distribution of the direct index, carried out in the second period, for relative optical air masses 1 and 2. The parameters α and β were estimated from Eqs. (8) and (9) using Eq. (24). $\alpha_2 = \beta_2 = 1$ was assumed for the second Beta probability function.

5. The frequency distribution of the diffuse index

The analysis of the frequency distribution of diffuse index (k_d) was carried out using a procedure similar to those applied before in sections 3 and 4. In this case the fitting was based on only one Beta probability function. Figure 5 shows the relative frequency distribution of k_d for four relative optical air masses. They indicate a positive asymmetry with similar characteristics for all relative optical air masses.

The relative frequency increases with relative optical air mass for $0.00 \leq k_d < 0.05$. This behavior is consistent and opposite to that found for k_{bh}, since both indices are complements of the clearness index.

The peak of the frequency distribution increases inversely proportional to the relative optical air mass, with a slight shift toward higher k_d values as the relative optical air mass increases. In the observed range of relative optical air mass, k_d ranges from 0.01 to approximately 0.81, while the maximum frequency ranges from 0.07 to 0.19. This implies that at noon-time and for a cloudless sky condition, the global solar irradiance at the surface is largest and, consequently, the averaged fraction of diffuse component reaches its minimum. On the other hand, the diffuse component reaches the maximum value at

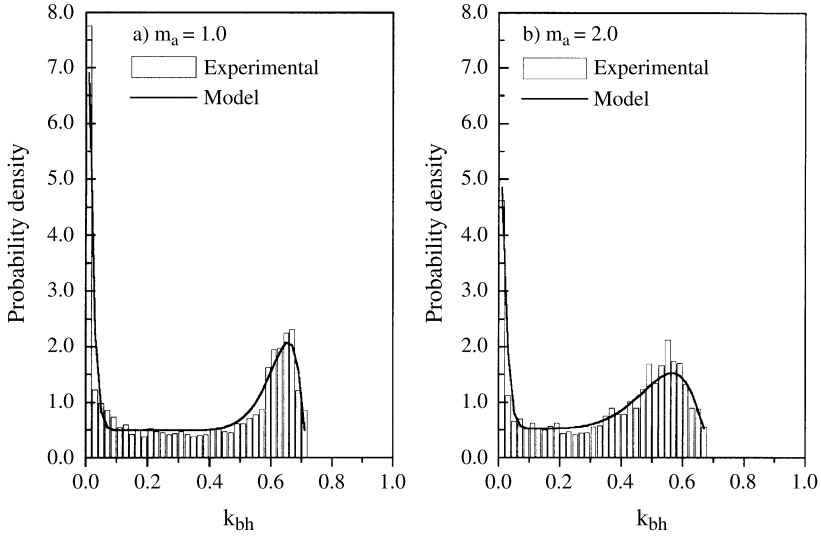

Fig. 4. Probability functions of k_{bh} modeled for **a**) $m_a = 1.0$ and **b**) $m_a = 2$. Observations were carried out in the second period (March 1999 to 2000)

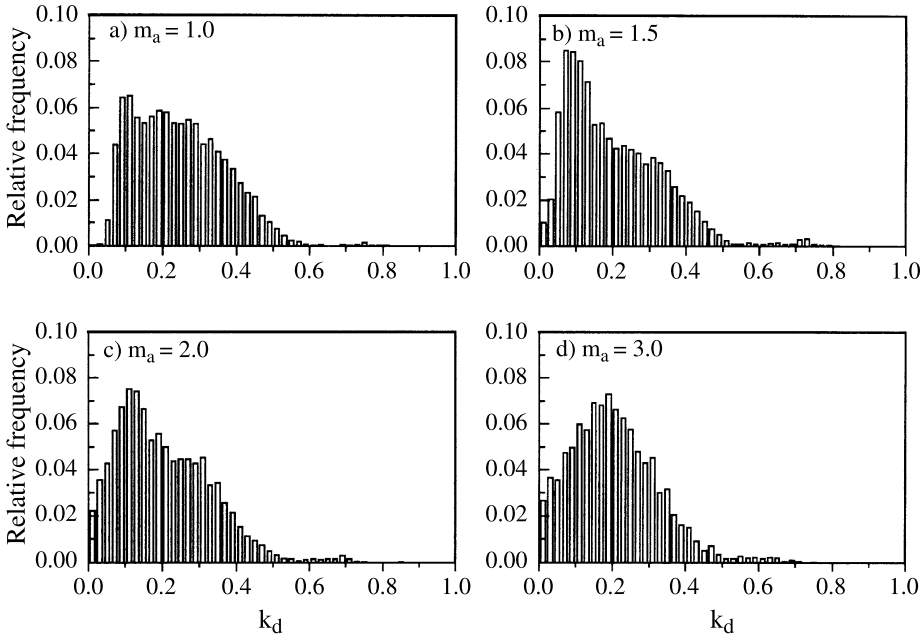

Fig. 5. Relative frequency distribution of k_d for **a**) $m_a = 1.0$, **b**) $m_a = 1.5$, **c**) $m_a = 2$ and **d**) $m_a = 3.0$. Observations carried out in the first period (March 1996 to 1999)

noon-time and for a cloudy sky condition. Only in the last case do the frequency distributions of k_d have a weak dependence on the relative optical air mass.

Table 7 presents the parameters of the experimental frequency distributions. The best fit for the frequency distribution parameter as a function of relative optical air mass was accomplished with the following polynomials:

$$\hat{\bar{k}}_d = 0.0084m_a^2 - 0.047m_a + 0.2708 \qquad (25)$$

$$\hat{s}^2 = -0.0007m_a^2 + 0.0039m_a + 0.0103 \qquad (26)$$

The accuracy of these expressions is equivalent to that of the polynomials derived in section 3.

Tovar et al. (1999) found an opposite trend for $\bar{k}_d$, which was attributed to the intensification of atmospheric scattering as the relative optical air mass increases. The resulting relative frequency distribution of k_d was not as sensitive to relative optical air mass variation as observed here. This divergence may be due to inherent climate spatial variation or even in the method of measuring diffuse solar radiation, because Tovar et al. (1999) measured diffuse solar radiation using a shadow-band. In our observations diffuse solar irradiance was calculated as difference of global minus horizontal direct solar irradiance measured with a pyrheliometer.

Table 7. Frequency distribution parameters of k_d for different values of relative optical air mass

m_a	$\bar{k}_d$	s^2
1.0	0.233	0.01344
1.5	0.217	0.01476
2.0	0.212	0.01523
3.0	0.205	0.01593

The validation of the model is illustrated in Fig. 6. The probability function was estimated from Eq. (7), solving Eqs. (25) and (26) in Eqs. (8) and (9).

6. Performance of the models

The degree of fit of each k_t curve can be objectively assessed using as criteria the statistical parameters MBE (Mean Bias Error) and RMSE (Root Mean Square Error) (Stone, 1993; Oliveira et al., 2002). The resulting MBE and RMSE are close to zero, indicating that the proposed functions perform satisfactorily. However, MBE = 0.0002 ($m_a = 2.0$) and MBE = 0.00005 ($m_a = 1.0$), also indicate that in all intervals of relative optical air mass considered here, the predicted values overestimate the observations.

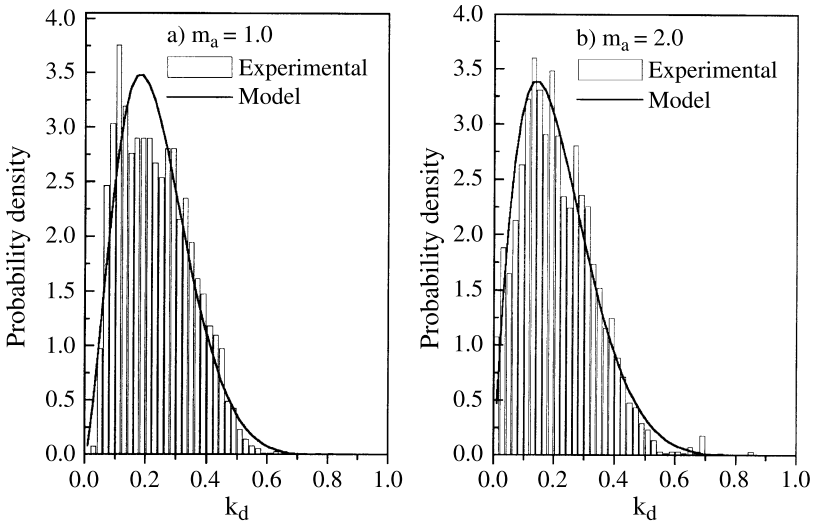

Fig. 6. Probability function of k_d modeled for **a**) $m_a = 1.0$ and **b**) $m_a = 2$. Observations were carried out in the second period (March 1999 to 2000)

A maximum dispersion is observed for relative optical air mass 1.0 (RMSE = 0.395). A minimum dispersion is found for relative optical air mass 1.5 (RMSE = 0.1675). A student t-test indicates that for all relative optical air mass intervals considered here, the proposed functions represent the observations at the 0.05 significance level (Oliveira et al., 2002). The coefficient of determination (R^2) 0.98 for all relative optical air masses confirms the quality of the functions proposed here.

The models for k_{bh} were tested against observations using values of MBE and RSME and the results were satisfactory. The models overestimate observations for all intervals of relative optical air mass, MBE = 0.0035 ($m_a = 3.0$) and MBE = 0.0121 ($m_a = 1.0$). Maximum dispersion occurred for relative optical air mass 1.0, with RMSE = 0.2962. Minimum dispersion occurred for relative optical air mass 2.0, with RMSE = 0.2405. The student t-test also indicates that the model reproduced the observations at the 0.05 significance level (Stone, 1993) for all relative optical air masses. The coefficient of determination (R^2) remained above 0.98 for all the conditions observed above.

Finally, the performance of the models for k_d, indicated by MBE and RMSE, are also satisfactory. For all values of relative optical air mass, the model overestimates the observed relative probability density distribution of k_d, MBE = 0.0035 ($m_a = 3.0$) and MBE = 0.0121 ($m_a = 1.0$). Maximum dispersion (RMSE = 0.2962) is observed at $m_a = 1.0$ and minimum (RMSE = 0.2405) at $m_a = 2.0$. The student t-test indicates that the models fit the experimental relative frequency distributions of k_d at the 0.05 significance level (Stone, 1993). In this case, the coefficients of determination (R^2) are also above 0.98 for all values of relative optical air mass.

7. Conclusion

The main goal of the present work was to model the probability distribution functions of clearness, beam and diffuse indices, in terms of relative optical air mass for short term variations of solar radiation at the surface, using five minute-averaged values, observed in the rural area of Botucatu, Brazil, during a four year period.

The effect of relative optical air mass on the frequency distribution of clearness, diffuse and direct indices of solar radiation can be summarized by the following conclusions:

- The relative frequency distributions of k_t are bimodal for all relative optical air mass between 1.0 and 3.0. The shape of these distributions is strongly dependent on the relative optical air mass. These characteristics can be explained in terms of the existence of two irradiance levels, corresponding to two extreme atmospheric conditions: totally clear and totally cloudy skies;
- The probability functions of k_t can be modeled by a combination of two Beta probability

functions. The parameters derived from the observed frequency distributions are strongly correlated to the relative optical air mass which can therefore be used to estimate α and β in the modeled probability functions of k_t.

- The probability functions of k_{bh} have an exponential shape between 0.01 and 0.05, regular form and an asymmetric characteristic. This distribution can be modelled by a combination of three Beta probability functions. The resulting model fits the observation very well;
- The observed frequency distribution of k_d is asymmetric and can be modelled with one Beta probability function. In this case, the model parameters have a weak dependence on the relative optical air mass.

Acknowledgment

The authors thank Fapesp and CNPq for the financial support of this research.

References

Berninger F (1994) Simulated irradiance and temperature estimates as a possible source of bias in the simulation of photosynthesis. Agric Forest Meteorol 71: 19–32

Gansler RA, Klein SA (1995) Beckman W. A. Investigation of minute solar radiation data. Solar Energy 55: 21–27

Iqbal M (1983) An introduction to solar radiation. Toronto, Canada: Academic Press, 390pp

Jurado M, Caridad JM, Ruiz V (1995) Statistical distribution of the clearness index with radiation data integrated over five minute intervals. Solar Energy 55: 469–473

Liu HYB, Jordan CR (1960) The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Solar Energy 4: 1–19

Skartveit A, Olseth JA (1992) The probability density and autocorrelation of short-term global and beam irradiance. Solar Energy 49: 477–487

Stone RJ (1993) Improved statistical procedure for the evaluation of solar radiation estimation models. Solar Energy 51: 289–291

Suehrcke H, McCormick PG (1988a) The frequency distribution of instantaneous insolation values. Solar Energy 40: 413–422

Suehrcke H, McCormick PG (1988b) The diffuse index of instantaneous solar radiation. Solar Energy 40: 423–430

Suehrcke H, McCormick PG (1989) Solar radiation utilizability. Solar Energy 43: 339–345

Tovar J, Olmo FJ, Alados-Arboledas L (1998) One minute global irradiance probability density distributions conditioned to the optical air mass. Solar Energy 62: 387–395

Tovar J, Olmo FJ, Batlles FJ, Alados-Arboledas L (1999) One minute k_b and k_d probability density distributions conditioned to the optical air mass. Solar Energy 65: 297–304

Oliveira AP, Escobedo JF, Machado AJ, Soares J (2002) Correlation models of diffuse solar-radiation applied to the city of São Paulo, Brazil. Applied Energy 71: 49–53

Authors' addresses: Dr. Hildeu Ferreira da Assunção (e-mail: hildeu@fca.unesp.br), João F. Escobedo (e-mail: escobedo@fca.unesp.br), Universidade Estadual Paulista, DRN – Facultade de Ciencias Agronomicas, P. O. Box 237, Botucatu, SP 18603-970, Brazil; Amauri P. Oliveira, Group of Micrometeorology, Department of Atmospheric Sciences, Institute of Astronomy, Geophysics and Atmospheric Sciences, University of São Paulo, São Paulo, Brazil.