Model to Generate Daily and Hourly Solar Radiation Sequences for Tropical Climates

Bao The Nguyen[*, **‡], Viet Van Hoang[*,**,***] and Hiep Chi Le[*,**]

[*]Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam

[**]Vietnam National University, Ho Chi Minh City, Vietnam

Address: 268 Ly Thuong Kiet street, District 10, Ho Chi Minh, Vietnam.

[***]Ly Tu Trong College, Ho Chi Minh City, Vietnam

Address: 390 Hoang Van Thu street, Tan Binh district, Ho Chi Minh, Vietnam.

(thebao@hcmut.edu.vn, hoangvanviethd@gmail.com, lechihiep@gmail.com)

‡ Corresponding Author: Bao The Nguyen; 268 Ly Thuong Kiet street, District 10, Ho Chi Minh, Vietnam.

Tel: (+84) 0906 33 11; Email: thebao@hcmut.edu.vn

Received: 10.10.2020 Accepted :19.11.2020

Abstract- The purpose of this study is to work out computational models to generate the sequences of daily and hourly solar radiation in tropical countries that are very important to calculate or simulate any solar thermal or electrical systems. In this study, the Aguiar's model is firstly used to generate daily clearness index series for Ho Chi Minh City and Da Nang, two cities presenting for two climate types in tropical region. Then a modified model of Aguiar is proposed to increase the accuracy in generating daily clearness index sequences for these two locations. In comparison with the results from original Aguiar's model, the modified model increases the accuracy of the mean and median of predicted daily clearness index sequences compared to those of measured data by 24.4% and 39.8% for Ho Chi Minh City and 34.8% and 47.2% for Da Nang, respectively. Then a modified model of Graham is suggested to predict hourly clearness index values for these two cities. The modified Graham's model increases the accuracy in predicting hourly solar radiation values about 2,5% in comparison with the accuracy of the original Graham's model. Especially, the model to generate the sequences of hourly k_t values proposed in this study is much simpler in comparison to the original model of Graham. Therefore, both proposed models in this work are expected to be used to generate daily and hourly solar radiation sequences for any locations in tropical regions because they do not need to use any location-dependent parameters.

Keywords Hourly clearness index; daily clearness index; monthly average daily clearness index; hourly solar radiation sequences; daily solar radiation sequences; Markov Transition Matrix; MTM library

1. Introduction

Computer simulation programs can be very useful tools in the process of designing solar energy systems. They enable users to optimize the performance of the solar systems in relation to various design parameters on the basic of the climatic and socio-economic conditions prevailing on the place of application. Consequently, they help the users to evaluate the cost effectiveness of the solar systems, to size the system components and hence to achieve not only the least cost, but also the most effective systems.

Weather data considered as forcing functions operating on the sets of equations used in the simulation programs to describe the performance of the solar systems. For design purposes, it is required to have at least one full year's data. Long term meteorology data provides the best estimates of representative behaviour of the weather for the location under investigation. According to Duffie and Beckman [1], weather records of at least 8 years are needed to provide a good representation when simulating solar heating systems. However, there are very few sites around the world where such long-term data are available.

An alternative solution to the use of long-term data is the use of typical meteorological year (TMY) data. The TMY concept is derived directly from the long-term data. Statistical parameters from the long-term data, such as the average, cross-correlations and distributions were determined for a number of different weather indices, for each month [2]. A hierarchy of criteria is then established to select 1 month from 23 years of data for each of the 26 locations. This approach has since been applied to create the TMY data for Canada and a similar test reference year (TRY) for Europe (Boland, 2008). Although this approach may reduce the computational effort in simulation studies of solar energy applications and reduce data sets, it is still based on the long-term data, which will not be available for most locations around the world, especially in developing countries.

As discussed by Nguyen and Hoang [3], most of developing countries are located in tropical regions. Unlike developed countries where measured solar radiation data are usually available, those of developing countries are often in shortages. For instance, among the 25 stations in a developing country like Vietnam where weather data is available, only twelve of them have records of both the sunshine duration and the global radiation data. The remainders only have sunshine duration records. Moreover, the solar radiation was measured at three-hour intervals rather than at every hour. Therefore, a computer model to generate hourly solar radiation data is needed for any simulation program in order to analyze the feasibility of the application of solar systems in tropical regions.

There are two methods used to overcome the lack of weather data: (i) using data from nearby stations or from similar climate locations (extrapolation method); and (ii) generating a weather data series (synthetic generation method). The first method is likely to result in significant errors [2]. Consequently, there have been numerous attempts at developing models using the second method.

Many researchers have developed stochastic simulation models of the sequences of weather data. Fernandez-Peruchena et al. [4] and Boland [2] used this approach to generate daily and hourly solar radiation values. Brecl and Topic [5] used a similar technique to generate daily and hourly solar radiation sequences for inclined surfaces from monthly average daily solar irradiation. Bright et al. [6] and Hofmann et al. [7] also used similar method to derive minutely irradiance time series from hourly data. Soubdhan and Emilion [8] even used the similar technique to generate secondly solar radiation sequences. Magnano et al. [9] applied stochastic function to generate synthetic sequences of half hourly temperature. One of the problems with most of these approaches is that the probability distribution functions (PDF) of the generated data are normal when stochastic models are used [10]. Gafurov et al. [11] incorporated spatial correlation of solar radiation (SCSR) into conventional stochastic solar radiation models, including Aguiar's, to generate monthly and daily solar irradiance time series.

Recently, some researchers have used different types of artificial neural network (ANN) to model values of total solar radiation on horizontal surfaces, such as [12], [13], [14], [15], [16]. The problem of these models is that they are "black boxes", and only mean values of daily global radiation have been analyzed, leading to no significant information can be obtained [10]. Mora-Lopez et al. [17] suggested using probabilistic finite automata (PFA) from machine learning theory to obtain values of daily total solar radiation. The drawback of this method is that the use of PFA is complicated and this approach has not been proved to be universally applied.

Another approach which has been widely used currently is the estimation of solar radiation irradiances from satellite data. Janjai et al. [18] developed a model to calculate global solar radiation from geostationary satellite then used it to generate hourly solar radiation maps of Thailand. A similar model has also been developed to build solar maps in Vietnam, both for global (or total) and direct (or beam) solar radiation components [19]. By using monthly average daily global irradiation data form NASA website, Berrizbeitia et al. [20] formulated the regression equations to estimate the monthly average hourly global and diffuse solar radiation values. However, one drawback of this approach is the use of other weather parameters such as ambient temperature and relative humidity [18], or cloud and rainfall [21]. These weather parameters are not always available in weather stations in developing countries [3]. Furthermore, the direct solar radiation data achieved from this approach is found to be incorrect in comparison with respectively measured data [3]. This is reinforced by Manju and Sandeep [22] when finding that surface measured data more suitable than satellite data in forecasting daily horizontal global solar radiation.

For the forecasting purpose, many models have been suggested as well. Wu and Chan [23] used a novel hybrid model of ARMA (Autoregressive Moving Average) and TDNN (Time Delay Neural Network) to predict hourly solar radiation in Singapore. Colak et al. [24] used ARMA and ARIMA (Autoregressive Integrated Moving Average) to predict hourly solar irradiance with one-period, two-period and three-period ahead and found the ARMA (1,2) and ARIMA (2,2,2) gave the best results. For daily solar radiation forecasting, Colak et al. [25] integrated the grey wolf optimizer to the multilayer perceptron algorithms and achieved the efficient prediction results. Some other hybrid methods have been employed such as wavelet-based recurrent neural networks [26], support vector machines [27], machine learning regressors [28], fuzzy optimization method [29], curve fitting methods [30], forecasting solar and air temperature together [31], [32], etc.

In this study, the stochastic technique is chosen to generate daily and hourly global solar radiation from monthly average daily clearness indices daily clearness index values for tropical countries. First, a stochastic model is used to generate synthetic of daily clearness index sequences from monthly average daily clearness indices daily clearness index values. Then the generated daily clearness index sequences are used to generate hourly clearness index series.

2. The Stochastic Model to Generate Daily Clearness Index Sequences

2.1. Model to generate daily solar radiation sequences in tropical climates

The daily radiation data for 300 months from nine stations having various types of climates was analysed by Aguiar et al. [33]. The authors discovered that the probability function for any given period seemed to have a form associated with only the average value of the daily clearness index, $\overline{K}_t$, for that period. Furthermore, they also found that any given daily radiation value had a significant correlation with only the radiation value of the preceding day in the sequence. Consequently, based on the MTM technique, Aguiar and his colleagues used the set of 300-month data to obtain ten 10 x 10 MTMs (called the MTM library). The model was tested by generating daily radiation sequences for many US locations, which had not been previously used in the derivation of the MTMs. This model gave better results in most cases when compared with the results from Graham's model [34], in terms of the statistic characteristics (e.g., average, variance and probability density function) and the sequential characteristics (e.g., the autocorrelation function). Another advantage of Aguiar's model lies in its computational simplicity.

For the daily radiation generation model, Aguiar's model has been chosen for the following reasons. First, among the nine locations that Aguiar used to develop the MTM library, two locations are represented for tropical regions, including one location, Ma Cau, has the mesothermal tropical forest climate (C_{aw}) and another location, Polana, (Mozambique) has type of tropical forest climate (A_w). Therefore, this approach would appear to be more suited as a model for tropical climates. Furthermore, because these nine locations consist of many different types of climate [33], this model is believed to be more universally applicable to any location. Secondly, Mustacchi et al. [35] found that the Markov chain model overcomes all other stochastic model approaches for simulation daily clearness (K_T) and hourly clearness (k_t) times series. This has been validated for many locations across Australia [36]. Fernandez-Peruchena et al. [4] also used Aguiar's model to generate hourly irradiation sequences for 3 locations in Spain and achieved the good agreements in comparison with measured data. Similar results were achieved by Brecl and Topic [5] when applying for locations in Slovenia. Secondly, the advantage of Aguiar's model is the ability to preserve probabilistic and sequential characteristics of measured data in generated series [11]. Thirdly, this approach is simple and reducing computing requirements by using the MTM library. For these reasons, the approach of Aguiar to develop a computer model for generating K_T sequences from monthly average values, $\overline{K}_T$ is chosen.

2.2. Data used to validate the model

In this study, global solar radiation was measured for one year in Ho Chi Minh City, representing for tropical forest or equational savannah climate (A_w) and one year in Danang City, representative for tropical or equational monsoon climate (A_m). A pyranometer was used to measure global solar flux every 5 minutes, continuously from 5.40 AM to 6.30 PM every day. Because the two above mentioned cities are located very near the Equator, the daylengths are changed much during the year and the maximum duration of the days is just 13 hours, from 5.40 AM to 6.30 PM. The data was recorded and stored automatically in a computer. Then, the hourly radiation I was achieved by averaging the solar flux of each hour. After that, daily global solar radiation on the horizontal surface H and monthly average daily global solar radiation on the horizontal surface $\overline{H}$ were obtained for two mentioned cities. Table 1 gives the values of monthly average daily global solar radiation on the horizontal surface $\overline{H}$ of Ho Chi Minh and Da Nang. The values of long-term measured average daily global solar radiation $\overline{H_{LT}}$, obtained from [19], are also given in this Table for comparison. As being shown, the errors between one-year measured data from this study and long-term measured ones are less than 10%, so the one-year measured data can be used to calculate the monthly average daily clearness $\overline{K_T}$ in this study.

Table 1. Monthly average daily global solar radiation on the horizontal surface (kWh/m^2) of Ho Chi Minh and Da Nang

		Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
HCM City	$\overline{H}$	3.61	4.99	5.04	5.20	4.65	4.83	4.80	4.90	4.43	4.07	4.16	3.85
	$\overline{H_{LT}}$	4.0	5.2	5.2	5.5	4.8	4.8	4.8	4.8	4.4	4.4	4.4	4.0
	%	-9.8%	-4.0%	-3.1%	-5.5%	-3.1%	0.6%	0.0%	2.1%	0.7%	-7.5%	-5.5%	-3.8%
Da Nang	$\overline{H}$	2.87	5.22	5.17	6.15	6.36	6.65	5.64	5.20	4.79	4.12	3.31	2.34
	$\overline{H_{LT}}$	3.0	5.6	5.7	6.0	6.1	6.2	6.0	5.0	4.8	4.2	3.4	2.5
	%	-4.3%	-6.8%	-9.3%	2.5%	4.3%	7.3%	-6.0%	4.0%	-0.2%	-1.9%	-2.6%	-6.4%

The 365 values of daily clearness index, KTmea., for each of the two above cities were calculated. The daily clearness index is defined as [1]:

$$K_{Tmea} = \frac{H}{H_0} \quad (1)$$

with H is the measured daily global solar radiation on the horizontal surface for Ho Chi Minh and Da Nang; H0 is the daily extraterrestrial radiation, given by:

$$HO = \frac{24}{\pi} G_{SC} 3600 \left\{ \left[1 + 0.033. \cos\left(\frac{360n}{365}\right) \right] \times \left[\cos\emptyset \cos\delta \sin\omega_S + \frac{\pi}{180} \omega_S \sin\emptyset \sin\delta \right] \right\} \quad (2)$$

In (2), GSC, n, ϕ, δ and ωs respectively are solar constance, day of the year, latitude of the location, declination angle and sunset hour angle, that are defined and formulated in [1].

The 12 values of monthly average daily clearness index, $\overline{K_T}$, for each of the two above cities were also obtained. The monthly average daily clearness index is defined as [1]:

$$\overline{k_T} = \frac{\bar{H}}{\overline{H_0}} \quad (3)$$

with $\bar{H}$ is the measured monthly average daily global solar radiation on the horizontal surface for Ho Chi Minh and Da Nang, given in Table 1; $\overline{H_0}$ is monthly average daily extraterrestrial radiation on a horizontal surface, calculated by using (2) with n is the average day of the months, defined and given in [1]. Table 2 gives the $\overline{K_T}$ values of the two investigated cities.

Table 2. Monthly average daily clearness index $\overline{K_T}$ of Ho Chi Minh and Da Nang

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
HCM City	0.42	0.53	0.50	0.50	0.45	0.47	0.47	0.47	0.44	0.42	0.47	0.46
Da Nang	0.36	0.59	0.53	0.59	0.60	0.63	0.53	0.50	0.48	0.45	0.40	0.30

2.3. The application of Aguiar's model

Based on the data on Table 2, Aguiar's model is used to generate 365 values of daily clearness index $K_{Tgen.}$ for Ho Chi Minh and Da Nang as follows. Firstly, the monthly average daily clearness index $\overline{K_T}$ of January is picked up (from Table 2). This $\overline{K_T}$ value will be correspondent to one matrix in MTM library and one line in Table "Min-Max". Secondly, the daily clearness index K_{Ti} of the previous day is used to choose the value of K_t^{min} and K_t^{max} on the line in Table "Min – Max" which is chosen in Step 1. The value of K_t^{min} and K_t^{max} will indicate the line on the matrix picked up in Step 1. Thirdly, a random number having value from 0 to 1 is chosen from a random number generator (available in Excel). Now, plus the values in the indicated line on the matrix from Step 2 until this sum equal and larger than the random number of Step 3. Then, the K_{Ti} value will be calculated by:

$$F_i(K_t) = \int_0^{K_t} P(K_t')dK_t' \quad (4)$$

In which, $F_i(K_t)$ is the distribution function for state i (the calculated date) and $P(K_t')$ is the values in the indicated line on the matrix from Step 2.

After the value of K_T for the first day of January is determined, the K_T value of the second day of January is computed by repeating all above mentioned steps, and so on, until 365 values of daily clearness index $K_{Tgen.}$ is generated.

Finally, the measured and generated values of are plotted in graphs of cumulative distribution function (CDF) and probability density function (PDF) for statistical comparison. Figures 1 and 2 graph the cumulative distribution function (CDF) of K_T for Ho Chi Minh and Da Nang respectively while Figures 3 and 4 plot the probability density function (PDF) of K_T for these cities.

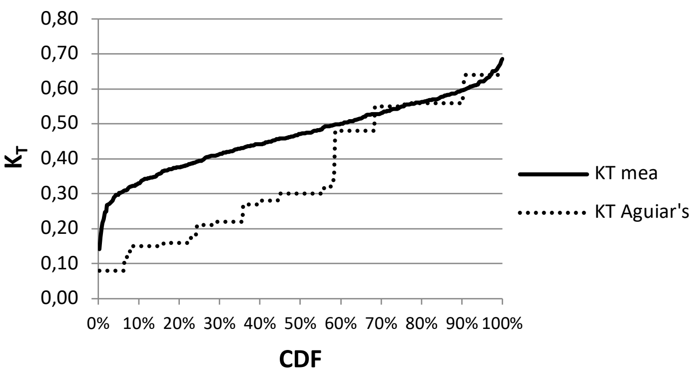

Fig. 1. Cumulative distribution function of K_T for Ho Chi Minh City

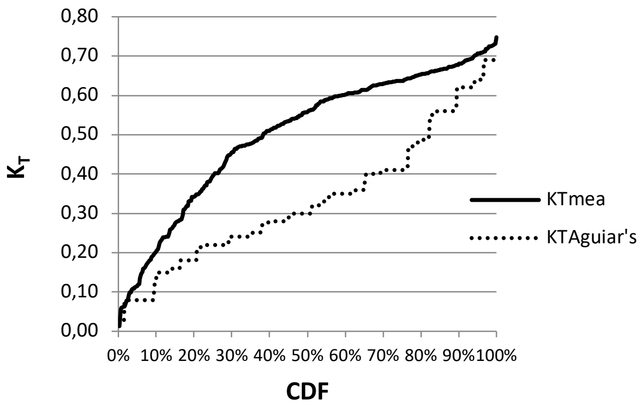

Fig. 2. Cumulative distribution function of K_T for Da Nang

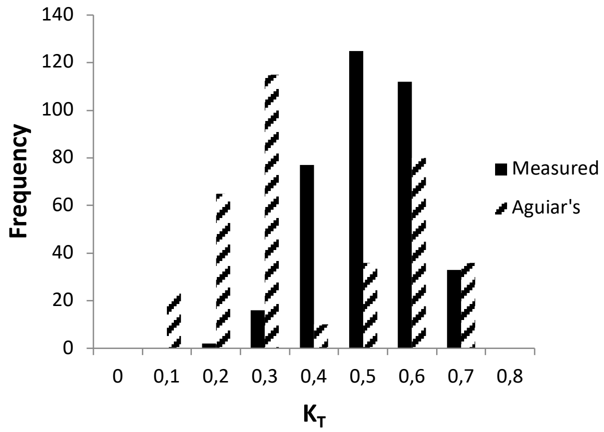

Fig. 3. Probability density function of K_T for Ho Chi Minh City

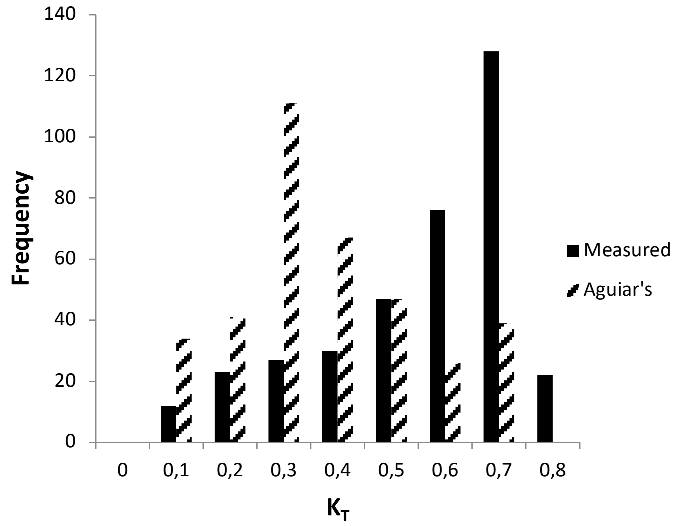

Fig. 4. Probability density function of K_T for Da Nang

Some statistic configurations, including mean, median, mean absolute error (MAE) and root mean square error (RMSE) of measured and generated K_T series of Ho Chi Minh City and Da Nang are also shown in Tables 3 & 4, respectively.

Table 3. Statistic configurations of K_T series of Ho Chi Minh City

	Mean	Median	Min	Max	Std.Dev.	MAE (%)	RMSE (%)
K_{Tmea}	0.47	0.47	0.14	0.69	0.10		
$K_{Tgen.}$	0.36	0.30	0.08	0.64	0.18		
Error (%)	26.5	44				50.1	24.6

Table 4. Statistic configurations of K_T series of Da Nang

	Mean	Median	Min	Max	Std.Dev.	MAE (%)	RMSE (%)
K_{Tmea}	0.50	0.56	0.01	0.75	0.18		
$K_{Tgen.}$	0.33	0.30	0.05	0.80	0.17		
Error (%)	41	60.5				23.7	28

As shown in Figures 1 to 4 and Tables 3 and 4, Aguiar's model seems to poorly generate daily radiation sequences for Vietnam. As shown in Tables 3 and 4, the error percentages in mean and median of generated sequences are 26.5% and 44% for Ho Chi Minh City and 41% and 60.5% for Da Nang, respectively, that are very high. These results are similar to the generated results of some locations that gave low confidence levels in comparison with the observed radiation sequences [33]. Then, Aguiar and his colleagues suggested that the Markov Transition Matrix (MTM) library should be modified to increase the accuracy of the model. Therefore, in this study, the MTM library has been modified as follows.

2.4. The modification of Aguiar's model

As mentioned above, Aguiar and his colleagues used the set of 300-month data and MTM technique to obtain the library of ten 10 x 10 MTMs. In this study, besides the data from Aguiar's research, obtained from [37], 24-month data of Ho Chi Minh and Da Nang were also used in order to build the MTMs. These MTMs were classified by different ranges of $\overline{K}_t$ values: one class for $\overline{K}_t \leq 0.30$, eight classes in steps of 0.05 between 0.30 and 0.70, and the last classes for $\overline{K}_t > 0.70$. With total of 324 months used in this study, the number of months in each class was 11, 4, 31, 51, 60, 66, 56, 19, 17 and 9, respectively. The procedure to construct Markov matrices is similar to Aguiar et al. [33], but with smaller intervals between states to build ten non-square 20 x 10 MTMs, as shown in Tables 5 to 14. In this tables, each row and column correspond to values of K_T between K_T^{min} and K_T^{max} for each state, which are given in Table 15.

Table 5. Modified Markov Transition Matrix for $\overline{K_T} \leq 0.30$

0.1145	0.1145	0.1665	0.1665	0.1040	0.1040	0.0210	0.0210	0.0415	0.0415	0.0210	0.0210	0.0210	0.0210	0.0105	0.0105	-	-	-	-
0.0835	0.0835	0.1595	0.1595	0.0970	0.0970	0.0695	0.0695	0.0485	0.0485	0.0140	0.0140	0.0210	0.0210	-	-	0.0070	0.0070	-	-
0.1250	0.1250	0.1250	0.1250	0.0450	0.0450	0.0680	0.0680	0.0455	0.0455	0.0230	0.0230	0.0230	0.0230	0.0115	0.0115	0.0340	0.0340	-	-
0.0790	0.0790	0.1185	0.1185	0.0790	0.0790	0.1315	0.1315	0.0130	0.0130	0.0265	0.0265	0.0395	0.0395	0.0130	0.0130	-	-	-	-
0.1055	0.1055	0.0265	0.0265	0.1055	0.1055	0.0790	0.0790	0.0260	0.0260	0.0260	0.0260	0.0790	0.0790	0.0525	0.0525	-	-	-	-
0.0625	0.0625	0.0625	0.0625	0.1250	0.1250	0.0940	0.0940	0.0315	0.0315	0.0620	0.0620	-	-	0.0625	0.0625	-	-	-	-
0.0200	0.0200	0.1200	0.1200	0.0400	0.0400	0.0600	0.0600	0.0400	0.0400	0.0400	0.0400	0.0600	0.0600	0.0600	0.0600	0.0400	0.0400	0.0200	0.0200
-	-	0.1250	0.1250	-	-	0.0625	0.0625	-	-	0.0620	0.0620	0.0625	0.0625	0.1250	0.1250	0.0315	0.0315	0.0315	0.0315
-	-	0.1250	0.1250	-	-	0.0625	0.0625	0.1250	0.1250	-	-	0.1250	0.1250	-	-	-	-	0.0625	0.0625
-	-	-	-	-	-	-	-	-	-	-	-	0.2500	0.2500	0.1250	0.1250	-	-	0.1250	0.1250

Table 6. Modified Markov Transition Matrix for $0.30 < \overline{K_T} \leq 0.35$

-	-	-	-	0.0455	0.0455	-	-	0.1820	0.1820	0.0450	0.0450	0.0910	0.0910	-	-	0.1365	0.1365	-	-
0.0590	0.0590	0.0590	0.0590	0.0880	0.0880	0.0590	0.0590	0.0295	0.0295	0.0590	0.0590	0.0880	0.0880	0.0290	0.0290	0.0295	0.0295	-	-
0.0335	0.0335	0.1335	0.1335	0.0335	0.0335	0.1000	0.1000	0.0330	0.0330	-	-	0.0665	0.0665	0.0665	0.0665	-	-	0.0335	0.0335
0.0590	0.0590	0.1175	0.1175	-	-	0.1175	0.1175	0.0295	0.0295	0.0880	0.0880	0.0590	0.0590	-	-	0.0295	0.0295	-	-
0.0385	0.0385	0.0770	0.0770	0.1540	0.1540	0.0385	0.0385	0.0770	0.0770	0.0380	0.0380	-	-	0.0385	0.0385	0.0385	0.0385	-	-
0.0415	0.0415	-	-	0.0835	0.0835	0.1250	0.1250	0.0415	0.0415	0.0835	0.0835	-	-	0.0415	0.0415	0.0835	0.0835	-	-
0.1110	0.1110	0.1110	0.1110	-	-	0.0555	0.0555	0.0560	0.0560	-	-	0.0555	0.0555	0.1110	0.1110	-	-	-	-
0.0455	0.0455	0.0910	0.0910	0.1365	0.1365	-	-	0.0450	0.0450	0.1365	0.1365	-	-	0.0455	0.0455	-	-	-	-
0.0555	0.0555	0.0555	0.0555	0.0560	0.0560	0.1110	0.1110	-	-	-	-	-	-	0.1110	0.1110	0.0555	0.0555	0.0555	0.0555
-	-	-	-	-	-	-	-	-	-	-	-	0.2500	0.2500	-	-	-	-	0.2500	0.2500

Table 7. Modified Markov Transition Matrix for $0.35 < \overline{K_T} \leq 0.40$

0.1035	0.1035	0.0440	0.0440	0.0885	0.0885	0.0880	0.0880	0.0440	0.0440	0.0145	0.0145	0.0885	0.0885	0.0145	0.0145	0.0145	0.0145	-	-
0.0600	0.0600	0.0500	0.0500	0.0700	0.0700	0.0800	0.0800	0.0600	0.0600	0.1100	0.1100	0.0500	0.0500	-	-	0.0100	0.0100	0.0100	0.0100
0.0385	0.0385	0.0615	0.0615	0.0925	0.0925	0.0615	0.0615	0.0385	0.0385	0.0695	0.0695	0.0460	0.0460	0.0615	0.0615	0.0305	0.0305	-	-
0.0240	0.0240	0.0560	0.0560	0.0475	0.0475	0.1030	0.1030	0.1030	0.1030	0.0950	0.0950	0.0475	0.0475	0.0240	0.0240	-	-	-	-
0.0295	0.0295	0.0680	0.0680	0.0590	0.0590	0.0685	0.0685	0.0490	0.0490	0.0590	0.0590	0.0590	0.0590	0.0785	0.0785	0.0295	0.0295	-	-
0.0070	0.0070	0.0480	0.0480	0.0695	0.0695	0.0765	0.0765	0.0625	0.0625	0.0695	0.0695	0.1040	0.1040	0.0280	0.0280	0.0210	0.0210	0.0140	0.0140
0.0365	0.0365	0.0505	0.0505	0.0580	0.0580	0.0725	0.0725	0.0435	0.0435	0.0795	0.0795	0.1015	0.1015	0.0435	0.0435	0.0145	0.0145	-	-
0.0095	0.0095	0.0185	0.0185	0.0555	0.0555	0.0280	0.0280	0.0370	0.0370	0.0555	0.0555	0.0925	0.0925	0.1480	0.1480	0.0370	0.0370	0.0185	0.0185
0.0175	0.0175	0.0340	0.0340	0.0175	0.0175	-	-	0.0175	0.0175	0.0515	0.0515	0.0860	0.0860	0.0690	0.0690	0.1895	0.1895	0.0175	0.0175
-	-	0.0830	0.0830	0.0835	0.0835	-	-	0.0835	0.0835	-	-	-	-	0.1665	0.1665	-	-	0.0835	0.0835

Table 8. Modified Markov Transition Matrix for $0.40 < \overline{K_T} \leq 0.45$

0.0835	0.0835	0.0830	0.0830	0.0835	0.0835	-	-	0.0415	0.0415	0.0625	0.0625	-	-	0.0835	0.0835	0.0625	0.0625	-	-
0.0580	0.0580	0.0580	0.0580	0.0750	0.0750	0.0585	0.0585	0.0415	0.0415	0.0585	0.0585	0.1000	0.1000	0.0335	0.0335	0.0085	0.0085	0.0085	0.0085
0.0245	0.0245	0.0425	0.0425	0.0670	0.0670	0.0790	0.0790	0.0490	0.0490	0.0550	0.0550	0.0670	0.0670	0.0670	0.0670	0.0305	0.0305	0.0185	0.0185
0.0190	0.0190	0.0450	0.0450	0.0700	0.0700	0.0705	0.0705	0.0835	0.0835	0.0705	0.0705	0.0450	0.0450	0.0705	0.0705	0.0195	0.0195	0.0065	0.0065
0.0045	0.0045	0.0690	0.0690	0.0370	0.0370	0.0465	0.0465	0.0970	0.0970	0.0695	0.0695	0.0835	0.0835	0.0465	0.0465	0.0370	0.0370	0.0095	0.0095
0.0180	0.0180	0.0095	0.0095	0.0585	0.0585	0.0495	0.0495	0.0720	0.0720	0.0900	0.0900	0.0900	0.0900	0.0585	0.0585	0.0360	0.0360	0.0180	0.0180
-	-	0.0230	0.0230	0.0305	0.0305	0.0305	0.0305	0.0680	0.0680	0.0795	0.0795	0.1365	0.1365	0.0835	0.0835	0.0490	0.0490	-	-
0.0080	0.0080	0.0280	0.0280	0.0400	0.0400	0.0640	0.0640	0.0520	0.0520	0.0400	0.0400	0.0800	0.0800	0.1040	0.1040	0.0680	0.0680	0.0160	0.0160
0.0055	0.0055	0.0265	0.0265	0.0105	0.0105	0.0215	0.0215	0.0640	0.0640	0.0480	0.0480	0.0370	0.0370	0.1115	0.1115	0.1385	0.1385	0.0370	0.0370
-	-	0.0375	0.0375	0.0185	0.0185	-	-	0.0370	0.0370	0.0370	0.0370	0.0370	0.0370	0.0370	0.0370	0.1665	0.1665	0.1295	0.1295

Table 9. Modified Markov Transition Matrix for $0.45 < \overline{K_T} \leq 0.50$

0.0600	0.0600	0.1000	0.1000	0.0800	0.0800	0.0600	0.0600	0.0600	0.0600	0.0600	0.0600	0.0400	0.0400	-	-	0.0200	0.0200	0.0200	0.0200
0.0500	0.0500	0.0400	0.0400	0.0600	0.0600	0.0700	0.0700	0.0700	0.0700	0.1000	0.1000	0.0900	0.0900	0.0200	0.0200	-	-	-	-
0.0225	0.0225	0.0570	0.0570	0.0335	0.0335	0.0855	0.0855	0.0625	0.0625	0.0855	0.0855	0.0400	0.0400	0.0795	0.0795	0.0285	0.0285	0.0055	0.0055
0.0075	0.0075	0.0300	0.0300	0.0420	0.0420	0.0495	0.0495	0.0955	0.0955	0.0765	0.0765	0.0765	0.0765	0.0575	0.0575	0.0575	0.0575	0.0075	0.0075
0.0120	0.0120	0.0150	0.0150	0.0490	0.0490	0.0490	0.0490	0.0825	0.0825	0.0975	0.0975	0.0975	0.0975	0.0700	0.0700	0.0215	0.0215	0.0060	0.0060
0.0075	0.0075	0.0130	0.0130	0.0310	0.0310	0.0620	0.0620	0.0720	0.0720	0.0850	0.0850	0.0850	0.0850	0.1110	0.1110	0.0310	0.0310	0.0025	0.0025
-	-	0.0065	0.0065	0.0225	0.0225	0.0540	0.0540	0.0560	0.0560	0.0875	0.0875	0.0940	0.0940	0.1120	0.1120	0.0585	0.0585	0.0090	0.0090
0.0040	0.0040	0.0115	0.0115	0.0270	0.0270	0.0330	0.0330	0.0465	0.0465	0.0625	0.0625	0.0955	0.0955	0.1265	0.1265	0.0915	0.0915	0.0020	0.0020
0.0030	0.0030	0.0115	0.0115	0.0305	0.0305	0.0165	0.0165	0.0335	0.0335	0.0415	0.0415	0.0695	0.0695	0.1110	0.1110	0.1610	0.1610	0.0220	0.0220
-	-	0.0225	0.0225	0.0450	0.0450	0.0455	0.0455	0.0230	0.0230	0.0230	0.0230	0.0680	0.0680	0.0455	0.0455	0.1365	0.1365	0.0910	0.0910

Table 10. Modified Markov Transition Matrix for $0.50 < \overline{K_T} \leq 0.55$

0.1250	0.1250	0.0895	0.0895	0.0535	0.0535	0.0535	0.0535	0.0715	0.0715	0.0355	0.0355	0.0535	0.0535	0.0180	0.0180	-	-	-	-
0.0665	0.0665	0.0110	0.0110	0.0445	0.0445	0.0555	0.0555	0.0780	0.0780	0.0890	0.0890	0.0555	0.0555	0.0665	0.0665	0.0335	0.0335	-	-
0.0315	0.0315	0.0240	0.0240	0.0715	0.0715	0.0240	0.0240	0.0875	0.0875	0.0715	0.0715	0.1030	0.1030	0.0475	0.0475	0.0395	0.0395	-	-
-	-	0.0110	0.0110	0.0390	0.0390	0.0555	0.0555	0.0780	0.0780	0.0780	0.0780	0.1220	0.1220	0.0835	0.0835	0.0220	0.0220	0.0110	0.0110
0.0080	0.0080	0.0140	0.0140	0.0190	0.0190	0.0345	0.0345	0.0800	0.0800	0.1095	0.1095	0.1150	0.1150	0.0800	0.0800	0.0375	0.0375	0.0025	0.0025
0.0065	0.0065	0.0125	0.0125	0.0150	0.0150	0.0465	0.0465	0.0720	0.0720	0.1010	0.1010	0.1075	0.1075	0.1095	0.1095	0.0275	0.0275	0.0020	0.0020
0.0030	0.0030	0.0205	0.0205	0.0175	0.0175	0.0320	0.0320	0.0450	0.0450	0.0900	0.0900	0.1685	0.1685	0.0960	0.0960	0.0245	0.0245	0.0030	0.0030
0.0055	0.0055	0.0105	0.0105	0.0145	0.0145	0.0175	0.0175	0.0660	0.0660	0.0615	0.0615	0.0920	0.0920	0.1855	0.1855	0.0410	0.0410	0.0060	0.0060
0.0040	0.0040	0.0080	0.0080	0.0080	0.0080	0.0120	0.0120	0.0355	0.0355	0.0515	0.0515	0.0795	0.0795	0.1350	0.1350	0.1545	0.1545	0.0120	0.0120
-	-	-	-	-	-	-	-	0.0295	0.0295	-	-	0.0295	0.0295	0.1470	0.1470	0.2060	0.2060	0.0880	0.0880

Table11. Modified Markov Transition Matrix for $0.55 < \overline{K_T} \leq 0.60$

0.1090	0.1090	0.0435	0.0435	-	-	0.0870	0.0870	0.0650	0.0650	0.0435	0.0435	0.0435	0.0435	0.0650	0.0650	0.0435	0.0435	-	-
0.0125	0.0125	0.0395	0.0395	0.0660	0.0660	0.0395	0.0395	0.0130	0.0130	0.0790	0.0790	0.0790	0.0790	0.0660	0.0660	0.0790	0.0790	0.0265	0.0265
0.0100	0.0100	0.0100	0.0100	0.0100	0.0100	0.0200	0.0200	0.0800	0.0800	0.0900	0.0900	0.0800	0.0800	0.1000	0.1000	0.0500	0.0500	0.0500	0.0500
0.0125	0.0125	0.0065	0.0065	0.0190	0.0190	0.0380	0.0380	0.0380	0.0380	0.0695	0.0695	0.0695	0.0695	0.1330	0.1330	0.1075	0.1075	0.0065	0.0065
0.0155	0.0155	0.0150	0.0150	0.0250	0.0250	0.0100	0.0100	0.0455	0.0455	0.0655	0.0655	0.0810	0.0810	0.1415	0.1415	0.0655	0.0655	0.0355	0.0355
0.0030	0.0030	0.0030	0.0030	0.0065	0.0065	0.0285	0.0285	0.0285	0.0285	0.0605	0.0605	0.1020	0.1020	0.1435	0.1435	0.0925	0.0925	0.0320	0.0320
0.0020	0.0020	0.0130	0.0130	0.0180	0.0180	0.0150	0.0150	0.0465	0.0465	0.0535	0.0535	0.0965	0.0965	0.1535	0.1535	0.0835	0.0835	0.0185	0.0185
0.0050	0.0050	0.0045	0.0045	0.0070	0.0070	0.0210	0.0210	0.0205	0.0205	0.0350	0.0350	0.0760	0.0760	0.2090	0.2090	0.1015	0.1015	0.0205	0.0205
0.0060	0.0060	0.0105	0.0105	0.0110	0.0110	0.0190	0.0190	0.0095	0.0095	0.0250	0.0250	0.0565	0.0565	0.1405	0.1405	0.1800	0.1800	0.0420	0.0420
0.0040	0.0040	0.0120	0.0120	0.0195	0.0195	0.0195	0.0195	0.0315	0.0315	0.0195	0.0195	0.0590	0.0590	0.0590	0.0590	0.1420	0.1420	0.1340	0.1340

Table 12. Modified Markov Transition Matrix for $0.60 < \overline{K_T} \leq 0.65$

0.0335	0.0335	0.0665	0.0665	0.0665	0.0665	0.0335	0.0335	0.0335	0.0335	0.1000	0.1000	0.0665	0.0665	0.0665	0.0665	0.0335	0.0335	-	-
0.0585	0.0585	0.0290	0.0290	0.0295	0.0295	0.0295	0.0295	0.0295	0.0295	0.0590	0.0590	0.0590	0.0590	0.1175	0.1175	0.0590	0.0590	0.0295	0.0295
-	-	0.0125	0.0125	0.0120	0.0120	0.0245	0.0245	0.0730	0.0730	0.0365	0.0365	0.0975	0.0975	0.1220	0.1220	0.0975	0.0975	0.0245	0.0245
0.0130	0.0130	-	-	0.0135	0.0135	0.0130	0.0130	0.0265	0.0265	0.0920	0.0920	0.1315	0.1315	0.0920	0.0920	0.1185	0.1185	-	-
0.0070	0.0070	-	-	0.0205	0.0205	0.0280	0.0280	0.0345	0.0345	0.0485	0.0485	0.0695	0.0695	0.1530	0.1530	0.1390	0.1390	-	-
0.0045	0.0045	0.0045	0.0045	0.0255	0.0255	0.0345	0.0345	0.0260	0.0260	0.0560	0.0560	0.1075	0.1075	0.1425	0.1425	0.0690	0.0690	0.0300	0.0300
0.0045	0.0045	0.0045	0.0045	0.0130	0.0130	0.0085	0.0085	0.0470	0.0470	0.0495	0.0495	0.1160	0.1160	0.1415	0.1415	0.1050	0.1050	0.0105	0.0105
0.0055	0.0055	0.0070	0.0070	0.0080	0.0080	0.0095	0.0095	0.0135	0.0135	0.0310	0.0310	0.0815	0.0815	0.2335	0.2335	0.1010	0.1010	0.0095	0.0095
0.0025	0.0025	0.0035	0.0035	0.0155	0.0155	0.0085	0.0085	0.0165	0.0165	0.0250	0.0250	0.0430	0.0430	0.1260	0.1260	0.2345	0.2345	0.0250	0.0250
-	-	-	-	0.0080	0.0080	0.0230	0.0230	0.0155	0.0155	0.0230	0.0230	0.0385	0.0385	0.0615	0.0615	0.2230	0.2230	0.1075	0.1075

Table 13. Modified Markov Transition Matrix for $0.65 < \overline{K_T} \leq 0.70$

-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5000	0.5000	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5000	0.5000	-	-
-	-	-	-	-	-	-	-	-	-	-	-	0.1250	0.1250	0.1250	0.1250	0.2500	0.2500	-	-
-	-	-	-	-	-	-	-	0.1250	0.1250	-	-	-	-	0.1875	0.1875	0.1250	0.1250	0.0625	0.0625
-	-	-	-	-	-	0.0415	0.0415	-	-	0.0835	0.0835	0.0835	0.0835	0.1250	0.1250	0.1665	0.1665	-	-
-	-	-	-	0.0210	0.0210	0.0210	0.0210	0.0210	0.0210	0.0410	0.0410	0.0415	0.0415	0.1460	0.1460	0.1460	0.1460	0.0625	0.0625
-	-	-	-	0.0160	0.0160	-	-	-	-	0.0160	0.0160	0.0645	0.0645	0.1935	0.1935	0.1775	0.1775	0.0325	0.0325
-	-	-	-	-	-	0.0190	0.0190	0.0190	0.0190	0.0375	0.0375	0.0235	0.0235	0.1700	0.1700	0.2075	0.2075	0.0235	0.0235
0.0020	0.0020	0.0025	0.0025	0.0035	0.0035	0.0035	0.0035	0.0055	0.0055	0.0150	0.0150	0.0260	0.0260	0.0705	0.0705	0.3270	0.3270	0.0445	0.0445
-	-	-	-	-	-	-	-	0.0300	0.0300	0.0305	0.0305	0.0150	0.0150	0.0150	0.0150	0.1745	0.1745	0.2350	0.2350

Table 14. Modified Markov Transition Matrix for $\overline{K_T}$> 0.70

-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5000	0.5000	-	-
0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
-	-	-	-	-	-	0.1250	0.1250	-	-	-	-	-	-	0.2500	0.2500	0.1250	0.1250	-	-
-	-	-	-	0.0710	0.0710	0.0715	0.0715	-	-	0.0715	0.0715	0.0715	0.0715	0.2145	0.2145	-	-	-	-
-	-	-	-	-	-	0.1000	0.1000	-	-	-	-	0.1000	0.1000	0.2000	0.2000	0.1000	0.1000	-	-
-	-	-	-	-	-	-	-	-	-	-	-	0.1115	0.1115	0.2220	0.2220	0.1665	0.1665	-	-
-	-	-	-	-	-	-	-	0.0400	0.0400	0.0400	0.0400	0.0400	0.0400	0.2400	0.2400	0.1200	0.1200	0.0200	0.0200
-	-	-	-	0.0135	0.0135	0.0045	0.0045	0.0135	0.0135	0.0090	0.0090	0.0675	0.0675	0.2615	0.2615	0.1260	0.1260	0.0045	0.0045
-	-	-	-	-	-	0.0115	0.0115	-	-	0.0215	0.0215	0.0215	0.0215	0.1630	0.1630	0.2555	0.2555	0.0270	0.0270
-	-	-	-	-	-	0.0715	0.0715	-	-	-	-	-	-	0.0715	0.0715	0.3570	0.3570	-	-

Table 15. Maximum and minimum value of K_T for each of the 10 above classes defined in the average monthly value of K_T

0.031	0.058	0.051	0.052	0.028	0.053	0.044	0.085	0.010	0.319
0.065	0.090	0.086	0.087	0.067	0.093	0.083	0.123	0.052	0.346
0.098	0.122	0.121	0.122	0.106	0.133	0.121	0.161	0.093	0.374
0.132	0.153	0.156	0.157	0.145	0.173	0.160	0.199	0.135	0.401
0.166	0.185	0.191	0.192	0.184	0.214	0.199	0.237	0.176	0.428
0.200	0.217	0.227	0.227	0.223	0.254	0.238	0.275	0.218	0.456
0.233	0.249	0.262	0.262	0.262	0.294	0.276	0.313	0.260	0.483
0.267	0.281	0.297	0.297	0.301	0.334	0.315	0.351	0.301	0.510
0.301	0.312	0.332	0.332	0.340	0.374	0.354	0.389	0.343	0.537
0.334	0.344	0.367	0.367	0.379	0.414	0.392	0.427	0.384	0.565
0.368	0.376	0.402	0.403	0.418	0.455	0.431	0.466	0.426	0.592
0.402	0.408	0.437	0.438	0.456	0.495	0.470	0.504	0.468	0.619
0.435	0.440	0.472	0.473	0.495	0.535	0.508	0.542	0.509	0.647
0.469	0.471	0.507	0.508	0.534	0.575	0.547	0.580	0.551	0.674
0.503	0.503	0.542	0.543	0.573	0.615	0.586	0.618	0.592	0.701
0.537	0.535	0.578	0.578	0.612	0.655	0.625	0.656	0.634	0.729
0.570	0.567	0.613	0.613	0.651	0.695	0.663	0.694	0.676	0.756
0.604	0.599	0.648	0.648	0.690	0.736	0.702	0.732	0.717	0.783
0.638	0.630	0.683	0.683	0.729	0.776	0.741	0.770	0.759	0.810
0.671	0.662	0.718	0.718	0.768	0.816	0.779	0.808	0.800	0.838
0.705	0.694	0.753	0.753	0.807	0.856	0.818	0.846	0.842	0.865

After the modification, the modified MTM library is used to regenerate the daily K_T values for Ho Chi Minh City and Da Nang. Figures 5 and 6 graph the cumulative distribution function (CDF) of K_T generated by using the modified MTM library for Ho Chi Minh City and Da Nang respectively while Figures 7 and 8 plot the probability density function (PDF) of the new set of K_T values for these cities. For comparison, the CDF and PDF of the daily K_T values generated by using Aguiar's model are also plotted in these figures. Some statistic configurations, including mean, median, mean absolute error (MAE) and root mean square error (RMSE) of measured K_T series and K_T generated by the modified model of Ho Chi Minh City and Da Nang are also shown in Tables 16 & 17, respectively.

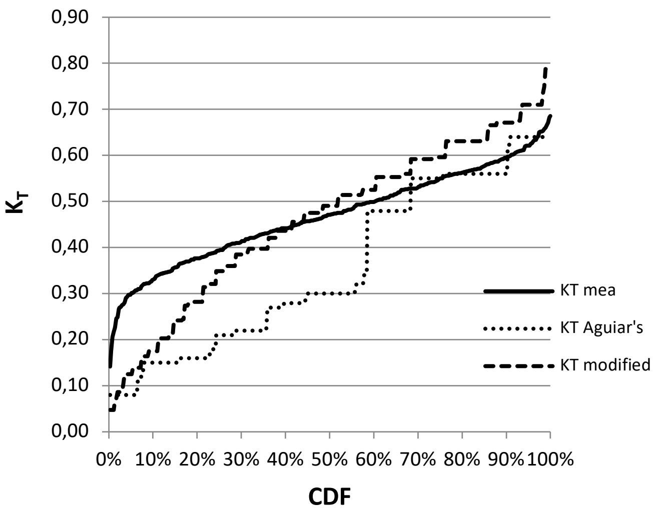

Fig. 5. Cumulative distribution function of three sets of K_T values for Ho Chi Minh City

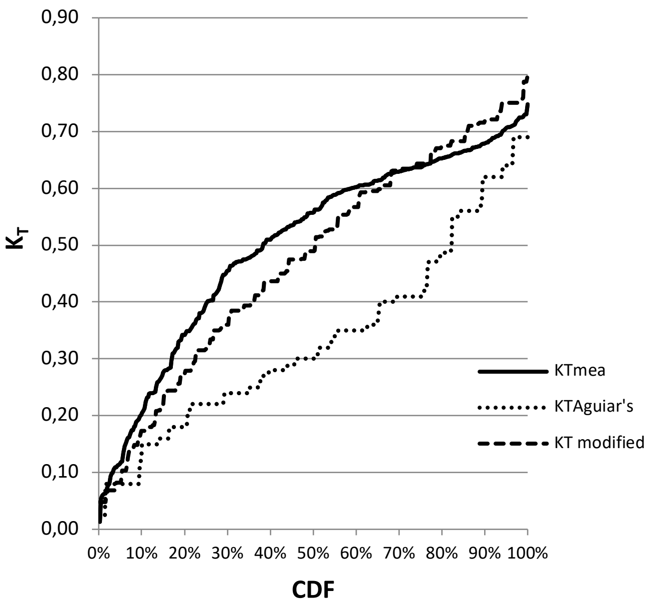

Fig. 6. Cumulative distribution function of three sets of K_T values for Da Nang

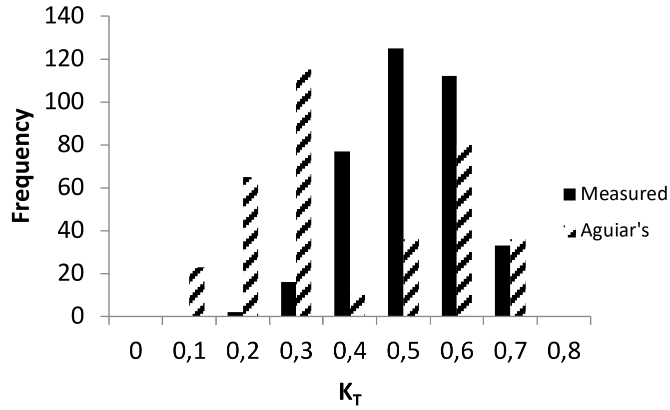

Fig. 7. Probability density function of three sets of K_T values for Ho Chi Minh City

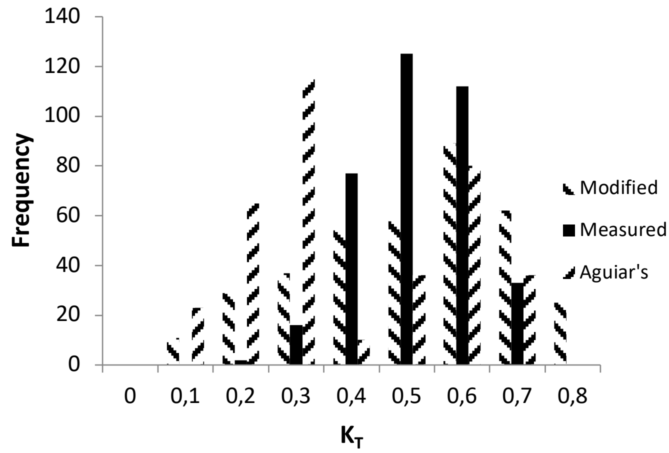

Fig. 8. Probability density function of three sets of K_T values for Da Nang

Table 16. Statistic configurations of K_T values of Ho Chi Minh City

	Mean	Median	Min	Max	Std.Dev.	MAE (%)	RMSE (%)
K_{Tmea}	0.47	0.47	0.14	0.69	0.10		
$K_{Tgen.}$	0.46	0.49	0.05	0.78	0.18		
Error (%)	2.1	4.2				16.8	8.1

Table 17. Statistic configurations of K_T values of Da Nang

	Mean	Median	Min	Max	Std.Dev.	MAE (%)	RMSE (%)
K_{Tmea}	0.50	0.56	0.01	0.75	0.18		
$K_{Tgen.}$	0.47	0.49	0.05	0.80	0.20		
Error (%)	6.2	13.3				4.5	5.1

Based on Figures 5 to 9 and Tables 16 & 17, it can be concluded that the modified model has given much better results compared to those from Aguiar's model. As shown in Tables 16 and 17, the error percentages in mean and median of generated sequences are 2.1% and 4.2% for Ho Chi Minh City and 6.2% and 13.3% for Da Nang, respectively. In comparison with the results from Aguiar's model, as shown in Tables 3 and 4 above, this modified model has increased the accuracy of predicted sequences by 24.4% and 39.8% for Ho Chi Minh City and 34.8% and 47.2% for Da Nang. Furthermore, this modified model is expected to be used to generate daily K_T series then daily radiation sequences for any locations in tropical regions because it does not need to use any location-dependent parameters.

3. MODEL TO GENERATE HOURLY CLEARNESS INDEX SEQUENCES IN TROPICAL CLIMATES

3.1. Model to generate hourly solar radiation sequences in tropical climates

As mentioned above, there have been many researches using stochastic models to generate daily and hourly solar radiation sequences ([2], [4], [5]). These researches were either based on Aguiar's approach [33] or Graham's approach [34].

For the stochastic model for the hourly radiation generation, Graham et al. [34] again used the hourly index clearness index k_t as variable rather than the hourly irradiation itself. Using the disaggregation technique, k_t is split into two components: a trend (or mean) component and a random component:

$$k_t = k_{tm} + \alpha \tag{5}$$

The trend component is then formulated as:

$$k_{tm} = \lambda + \rho \exp(-\kappa m) \tag{6}$$

where m is the air mass, taken at the midpoint of the hour. The parameters λ, ρ and κ are found to be unique functions of the daily clearness K_t:

$$\lambda(K_t) = K_t - 1.167K_t^3\ (1 - K_t) \tag{7a}$$

$$\rho(K_t) = 0.979(1 - K_t) \tag{7b}$$

$$\kappa(K_t) = 1.141(1 - K_t)/\ K_t \tag{7c}$$

The random component α, found for any hour event by $\alpha = k_t - k_{tm}$, is grouped for the days with approximately equal daily clearness K_t. It is found that the standard deviation σ_α of α varied from group to group. Furthermore, σ_α is strongly dependent on K_t, but is only a weak function of the zenith angle θ_z [34]. Therefore, σ_α may be expressed as:

$$\sigma_\alpha(K_t) = 0.16\sin(\pi K_t/0.90) \tag{8}$$

Because these random variables are non-Gaussian, they are first transformed into a Gaussian variables β, using the Gaussian mapping technique. The series of β was fitted using different stochastic models. The model with the best fit was the AR (1) model:

$$\beta_t = \phi\beta_{t-1} + \vartheta_t \tag{9}$$

where:

β_{t-1} is the value of the variable at t-1

ϕ is the coefficient, $\phi = 0.54$ is suggested for use in a universal model.

ϑ_t is a random number from a normal distribution with zero mean and a standard deviation $\sqrt{1-\phi}$

The values of β are then mapped into α values using the Beta distribution:

$$P(\alpha : K_t) = P(k_t : K_t) = \frac{\Gamma(p+q)u^{p-1}(1-u)^{q-1}}{\Gamma(p)\Gamma(q)(k_{tu}-k_{tl})} \quad (10)$$

in which: Γ is the gramma function; u is the random variable normalized from hourly clearness index kt variable within the (0,1) range; p and q are parameters evaluated by equating the data estimates of the mean and variance of u; k_{tu} and k_{tl} are upper limit and lower limit of k_t values, respectively [34].

In developing another hourly radiation model, Aguiar et al. [33] used an approach similar to Graham's approach but with some significant different. First, the authors discovered that hourly clearness index k_t was strongly dependent on both daily clearness K_t and the solar altitude angle h_s, and not just on the daily clearness K_t. Therefore, the standard deviation σ_α could be expressed as:

$$s_a (K_t, h_s) = A \exp [B (1 - \sin h_s)] \quad (11a)$$

with:

$$A = 0.14 \exp [-20.0(K_t - 0.35)^2] \quad (11b)$$

$$B = 3.0(K_t - 0.45)^2 + 16.0K_t^5 \quad (11c)$$

Secondly, instead of using an AR (1) model to fit the random component of the hourly clearness k_t which uses a constant coefficient ϕ, Aguiar et al. [33] suggested a K_t dependent equation for ϕ as:

$$\Phi = 0.38 + 0.06 \cos (7.4K_t - 2.5) \quad (12)$$

Thirdly, they introduced the clear-sky value k_{cs} to limit unreal generated k_t values. They also derived different equations to calculate the trend component k_{tm}. Finally, the authors included a normalization procedure to correct the nonstationary and time-homogeneous character of k_t rather than the Gaussian mapping technique, as proposed by Graham.

As mentioned above, Fernandez-Peruchena et al. [4] and Brecl and Topic [5] successfully used Aguiar's model to generate hourly solar irradiance in Spain and Slovenia, respectively whereas HOMER [38] uses Graham's approach to generate hourly solar radiation sequences in its software. Nguyen and Pryor [36] compared Aguiar's model and Graham in generation of hourly solar radiation for 6 locations in Australia and found that Graham's model generated better synthetic solar radiation sequences for all 6 locations. Therefore, in this study, Graham's model is chosen to generate hourly solar irradiances in tropical regions.

However, Graham's model has some drawbacks as follows:

- The use of Gaussian mapping technique to treat the random component of kt values, as briefly described above, is very complicated.
- The errors of hourly clearness index series generated from Graham's model compared with measured data are bigger than those from the modified model in this study that will be presented below.

To overcome to first drawback of Graham's model, in this study, the Gaussian values of β calculated by Equation (5) will be converted to non-Gaussian distribution h by the Norminv function in MATLAB. Then, the random components of k_t values are computed by:

$$\alpha = h \times \sigma_\alpha (K_t) \quad (13)$$

with the standard deviation $\sigma_\alpha(K_t)$ calculated from Equation (8).

This approach not only makes the procedure to generate hourly clearness index sequences much simpler compared to Graham's method, but also generate more accurate results. Table 18 presents the errors of generated hourly solar radiation sequences from Graham's model and the modified model of this study in comparison with measured solar irradiance sequences for Ho Chi Minh City in 20 times of running the generation program written by using MATLAB. As shown in the Table, the errors of the modified model of this study are always smaller than those of original Graham's model in 20 times of generation. Averagely, the modified model increases the accuracy in predicting hourly solar radiation values about 2,5% in comparison with the accuracy of the original Graham's model.

Table 18. The errors of generated hourly solar radiation compared to measured values in Ho Chi Minh City

Time of running the generation program	Error of generated hourly irradiances compared to measured values (%)	
	Graham's Model	**Modified Model of this study**
1	2.87%	0.74%
2	5.31%	2.70%

3	5.75%	2.49%
4	3.81%	1.46%
5	11.15%	9.47%
6	5.04%	2.84%
7	4.22%	2.06%
8	5.19%	3.60%
9	4.12%	2.27%
10	4.70%	2.18%
11	6.40%	4.40%
12	7.33%	5.29%
13	6.32%	4.22%
14	5.07%	2.60%
15	5.03%	2.78%
16	7.52%	4.67%
17	4.88%	2.52%
18	3.24%	1.48%
19	8.15%	6.44%
20	8.68%	5.36%

Therefore, in this study, the modified model of Graham's approach is used to generate hourly clearness index sequences. Figure 9 shows the procedure of this generation program.

The steps in Figure 9 as follows:

- Choose a month to calculate
- Choose the date of the month, from first date (i.e., 1) to the last date (i.e., max)
- Read the K_T value of that date from a file or from the result of program to calculate K_T value in Section 2. Then use Equation (8) to calculate σK_T
- Calculate k_{tm} by using Equation (6)
- Read a random number ε_t from Gaussian distribution
- Calculate $\chi(hr)$ by using the formula: $\chi(hr) = 0{,}54 \times \chi(hr-1) + \varepsilon_t$
- Calculate the cumulative distribution of function of a normally distributed variable Fnormal.
- Calculate kt by using the formula: $k_t = k_{tm} + F_{normal} \times \sigma K_T$

MATLAB is used to write the generation program for hourly kt sequences, then the hourly global solar radiation on the horizontal surface I values are computed by using Equation (12).

In Figure 9:

Φ is the latitude of the location

Lst and Lloc are the standard meridian for the local time zone and the longitude of the location, respectively.

j is month of the year (i.e., j = 1 is January, j = 2 is February, etc.)

i is the date of a month

ω_s is the sunset hour angle for the considered day

Kt [i][j] is the daily clearness index of i^{th} day in the j^{th} month

ω is the hour angle of the current hour.

k_{tm} is the "long-term" average kt values

σ_{kt} is the standard deviation of kt about the long-term average k_{tm} value

ε_t is a random number from a Gaussian distribution.

hr is the hour considered.

χ is a normally distributed stochastic variable with a mean of 0 and a variance of 1.

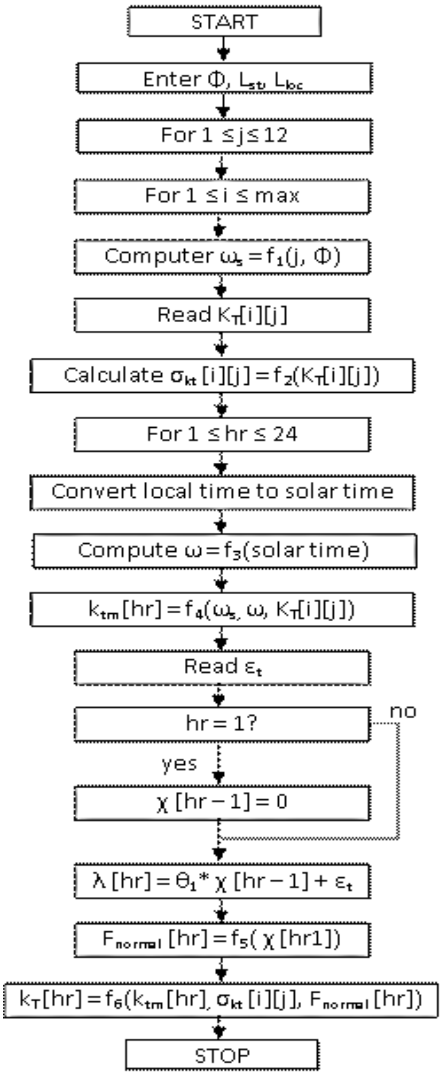

Fig. 9. The schematic diagram of the procedure to generate hourly kt sequences

θ_1 is the parameter of the first order autoregressive model.

F_{normal} is the function to convert a normally distributed variable to a non-Gaussian distributed variable.

3.2. Validating the generated hourly clearness index sequences in tropical region

The generated daily clearness index values are used as the input for the generation program of hourly k_t sequences in this study. Then, the generated k_t sequences are compared with the measured $k_{tmea.}$ Series, in which the measured hourly clearness index $k_{tmea.}$ is defined as:

$$k_{tmea.} = \frac{I}{I_0} \tag{14}$$

with I is the measured hourly global solar radiation on the horizontal surface for Ho Chi Minh and Danang; I0 is the hourly extraterrestrial radiation from hour angle ω1 to ω2, given by:

$$I0 = \frac{12}{\pi} G_{SC} \times 3600 \left\{ \left[1 + 0{,}033.\cos\left(\frac{360n}{365}\right) \right] \times \left[\cos\phi \cos\delta \,(\sin\omega_2 - \sin\omega_1) + \frac{\pi}{180}(\omega_2 - \omega_1) \sin\phi \sin\delta \right] \right\} \tag{15}$$

The cumulative distribution function (CDF) of hourly k_t for Ho Chi Minh and Da Nang are graphed in Figures 10 and 11 respectively while Figures 12 and 13 plot the probability density function (PDF) of hourly k_t for these cities.

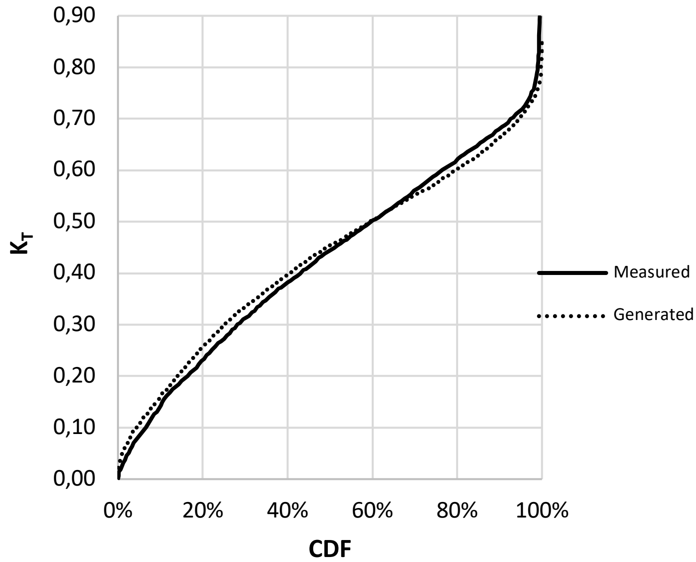

Fig. 10. Cumulative distribution function (CDF) of hourly k_t for Ho Chi Minh City

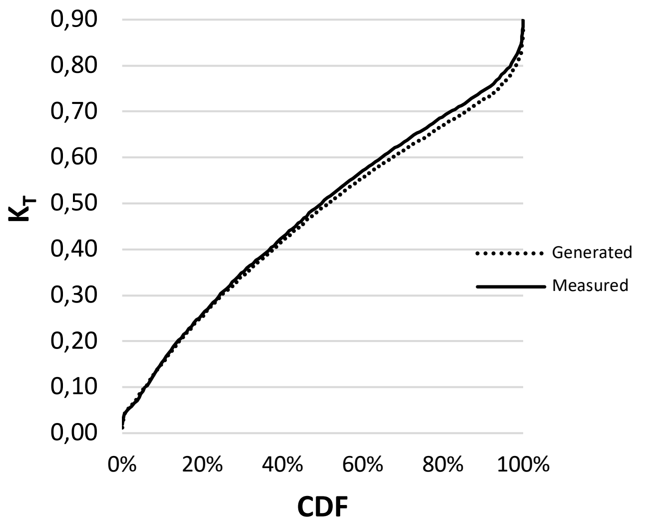

Fig. 11. Cumulative distribution function (CDF) of hourly k_t for Da Nang

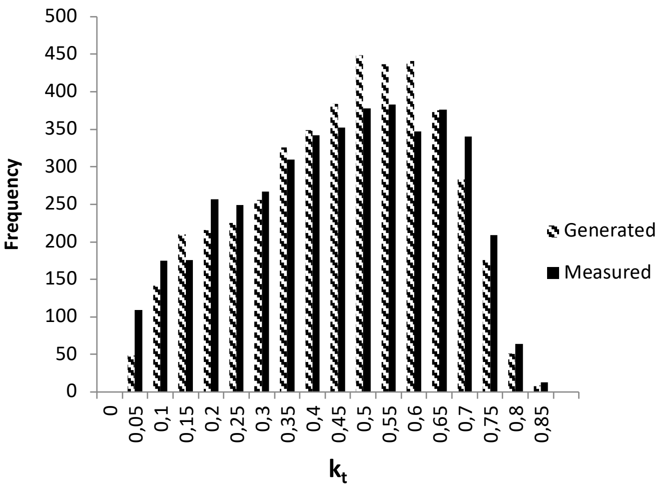

Fig. 12. Probability density function (PDF) of hourly k_t for Ho Chi Minh City

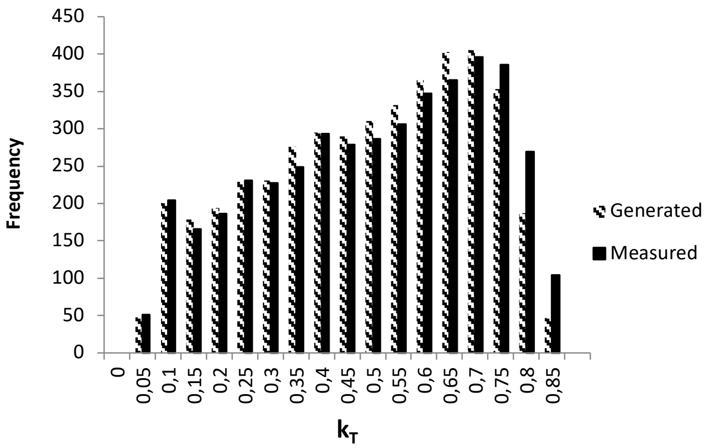

Fig. 13. Probability density function (PDF) of hourly k_t for Da Nang

Some statistic configurations, including mean, median, mean absolute error (MAE) and root mean square error (RMSE) of measured and generated k_t series of Ho Chi Minh City and Da Nang are also shown in Tables 19 & 20, respectively.

Table 19. Statistic configurations of hourly measured and generated k_t series of Ho Chi Minh City

	Mean	**Median**	**Min**	**Max**	**Std.Dev.**	**MAE (%)**	**RMSE (%)**
k_{tmea}	0.426	0.443	0.001	0.899	0.197		
$k_{tgen.}$	0.433	0.453	0.007	0.858	0.191		
Error (%)	3.0	4.8				2.0	0.03

Table 20. Statistic configurations of hourly measured and generated k_t series of Da Nang

	Mean	**Median**	**Min**	**Max**	**Std.Dev.**	**MAE (%)**	**RMSE (%)**
k_{tmea}	0.459	0.491	0.019	0.905	0.210		
$k_{tgen.}$	0.465	0.492	0.012	0.884	0.217		
Error (%)	2.5	0.6				3.3	0.03

As shown in the above Figures and Tables, the model in this study was successfully generated hourly k_t series for two cities representing for two tropical climate types with acceptable accuracies. As stochastic models have been approved to have universal characteristics, as mentioned above, the model in this study is expected to be applied for any locations in tropical regions.

4. Conclusion

In this study, the Aguiar's model was firstly used to generate daily clearness index series for Ho Chi Minh City and Da Nang, two cities presenting for two climate types in tropical region. Then a modified model of Aguiar was proposed to increase the accuracy in generating daily clearness index sequences for these two locations. In comparison with the results from original Aguiar's model, the modified model increased the accuracy of the mean and median of predicted daily clearness index sequences compared to those of measured data by 24.4% and 39.8% for Ho Chi Minh City and 34.8% and 47.2% for Da Nang, respectively. Then a modified model of Graham was suggested to predict hourly clearness index values for these two cities. The modified Graham's model increased the accuracy in predicting hourly solar radiation values about 2,5% in comparison with the accuracy of the original Graham's model. Especially, the model to generate the sequences of hourly k_t values proposed in this study is much simpler in comparison to the original model of Graham. Therefore, both proposed models in this work are expected to be used to generate daily and hourly solar radiation sequences for any locations in tropical regions because they do not need to use any location-dependent parameters.

Acknowledgement

This research is funded by Ho Chi Minh City University of Technology, VNU-HCM, under grant number 20190108.

References

[1] J. Duffie, and W. Beckman, Solar Engineering for Thermal Processes - 4th Edition. John Wiley & Sons, Inc., Hoboken, New Jersey, 2013.

[2] J. Boland, "Time series modelling of solar radiation". Modelling Solar Radiation at the Earth's Surface, V. Badescu (Ed.), Springer-Verlag Berlin Heidelberg, 2008, pp. 283-312.

[3] B. Nguyen, and V. Hoang, "The study of direct solar radiation data in the project of mapping the solar resource and potential in Vietnam". The 5th International Conference on Sustainable Energy 2017, Ho Chi Minh City, Vietnam.

[4] C. Fernandez-Peruchena, L. Ramirez, I. Pagola, and M. Gaston, "Assessment of models for estimation of hourly irradiation series from monthly mean values". 15th SolarPACES Conference, Berlin, Germany, pp.121-126. Hal-00919043.

[5] K. Brecl and M. Topic, "Development of a stochastic hourly solar irradiation model". International Journal of Photoenergy Volume 2014, DOI: 10.1155/2014/376504, 7 pages.

[6] J. Bright, C. Smith, P. Taylor, and R. Crook, "Stochastic generation of synthetic minutely irradiance time series derived from mean hourly weather observation data". Solar Energy 115, pp. 229-242.

[7] M. Hofmann, S. Riechelmann, C. Crisosto, R. Mubarak, and G. Seckmeyer, "Improved synthesis of global irradiance with one-minute resolution for PV system simulations". International Journal of Photoenergy Volume 2014, DOI: 10.1155/2014/808509. 10 pages.

[8] T. Soubdhan, and R. Emilion, "Stochastic differential equation for modelling global solar radiation sequences". Modelling, Identification and Control, GOSIER, France. DOI: 10.2316/P.2010.702-099. Hal-01823269.

[9] L. Magnano, J. Boland, and R. Hyndman, "Generation of synthetic sequences of half hourly temperature". Environmetrics, Vol. 19, No. 8, pp. 818-835.

[10] L. Mora-Lopez, "A new procedure to generate solar radiation time series from machine learning theory". Modelling Solar Radiation at the Earth's Surface, V. Badescu (Ed.), Springer-Verlag Berlin Heidelberg, 2008, pp. 313-326.

[11] T. Gafurov, J. Usaola, and M. Prodanovic, "Incorporating spatial correlation into stochastic generation of solar radiation data". Solar Energy 115, pp. 74-84.

[12] W. Hammed, B. Sawadi, S. Al-Kamil, M. Al-Radhi, and R. Abd-Alhameed, "Prediction of solar irradiance based on artificial neural networks". Invention Vol. 4, No. 45. DOI: 10.3390/inventions40300045.

[13] M. Vakili, S. Sabbagh-Yazdi, and K. Kalhor, "Using artificial neural networks for prediction of global solar radiation in Tehran considering particulate matter air pollution". Energy Procedia. Vol. 74, pp. 1205-1212.

[14] S. Priya, and M. Idbal, "Solar radiation prediction using artificial neural networks". International Journal of Computer Applications Vol. 116, No. 6.

[15] F. Tymvios, S. Michaelides, and C. Skouteli, "Estimation of surface solar radiation with artificial neural networks". Modelling Solar Radiation at the Earth's Surface, V. Badescu (Ed.), Springer-Verlag Berlin Heidelberg, 2008, pp. 221-246.

[16] A. Sfetsos, and A. Coonick, "Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques". Solar Energy 68, pp. 169-178.

[17] L. Mora-Lopez, J. Mora, R. Morales-Bueno, and M. Sidrach-de-Cardona, "Modelling time series of climatic parameters with probabilistic finite automata". Environmental Modelling & Software, Vol. 20, pp. 753-760.

[18] S. Janjai, P. Pankaew, and J. Laksanaboonsong, "A model for calculating hourly global solar radiation from satellite data in the tropics". Applied Energy, Vol. 86,

No. 9, DOI: 10.1016/j.apenergy.2009.02.005, pp. 1450–1457.

[19] CIEMAT, CENER & IDEA. Maps of Solar Resource and Potential in Vietnam. CIEMAT, CENER & IDAE with support from AECER in collaboration with GDE/MoIT, 2015.

[20] S. Berrizbeitia, E. Gago and T. Muneer, "Empirical Models for the Estimation of Solar Sky-Diffuse Radiation. A Review and Experimental Analysis". Energies, 13, 701; DOI:10.3390/en13030701, 23 pages

[21] D. Laslett, C. Creagh, and P. Jennings, "A method for generating synthetic hourly solar radiation data for any location in the south west of Western Australia, in a world wide web page". Renewable Energy, Vol. 68, DOI: 10.1016/j.renene.2014.01.015, pp. 87–102.

[22] S. Manju, M. and Sandeep, "Prediction and performance assessment of global solar radiation in Indian cities: A comparison of satellite and surface measured data". Journal of Cleaner Production, Vol. 230, pp. 116-128.

[23] J. Wu, and C. Chan, "Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN". Solar Energy, vol. 85, pp. 808-817.

[24] I. Colak, M. Yesilbudak, N. Genc, and R. Bayindir, "Multi-period prediction of solar radiation using ARMA and ARIMA models". 2015 IEEE 14th Conference on Machine Learning and Applications, DOI: 10.1109/ICMLA.2015.33, pp. 1045-1049.

[25] M. Colak, M. Yesilbudak, and R. Bayindir, "Forecasting of daily total horizontal solar radiation using grey wolf optimizer and multilayer perceptron algorithms". International Conference on Renewable Energy Research and Application, pp. 939-942. Brasov, Romania.

[26] G. Capizzi, C. Napoli, and F. Bonanno, "Innovative second-generation wavelets construction with recurrent neural networks for solar radiation forecasting". IEEE Transactions on Neural Networks and Learning System, vol 23, No.11, pp. 1805-1815.

[27] J. Chen, H. Liu, W. Wu and D. Xie, "Estimation of monthly solar radiation from measured temperatures using support vector machine – A case study". Renewable Energy, Vol. 36, No. 1, pp. 413-420.

[28] L. Cornejo-Bueno, C. Casanova-Mateo, J. Sanz-Justo and S. Salcedo-Sanz, "Machine learning regressors for solar radiation estimation from satellite data". Solar Energy, 183, pp. 768-775.

[29] H. Othmani, F. Sassi, D. Mezghani and A. Mami, "Fuzzy optimization of a photovoltaic pumping system: Implementation and measurements", International Journal of Renewable Energy Research, vol. 7, No. 3, pp. 1115-1124, 2017.

[30] M. Yesilbudak, M. Colak, R. Bayindir and H. Bulbul, "Very-short term modelling of global solar radiation and air temperature data using curve fitting methods". IEEE 6th International Conference on Renewable Energy Research and Applications, pp. 1144-1148, 5-8 November 2017, CA, USA.

[31] N. Genc, M. Yesilbudak and M. Colak, "A case study on the investigation of solar regime in Van, Turkey", IEEE 4th International Conference on Renewable Energy Research and Applications, pp. 954-958, 22-25 November 2015, Palermo, Italy.

[32] R. Al-Hajj, A. Assi and F. Batch, "An evolutionary computing approach for estimating global solar radiation", IEEE 5th International Conference on Renewable Energy Research and Applications, pp. 367-378, 20-23 November 2016, Birmingham, UK.

[33] R. Aguiar, M. Collares-Pereira, and J. Conde, "Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices". Solar Energy 40, pp, 269-279.

[34] V. Graham and K. Hollands, "A method to generate synthetic hourly solar radiation globally". Solar Energy, Vol. 44, No. 6, pp. 333-341.

[35] C. Mustacchi, V. Cena, and M. Rocchi, M. "Stochastic simulation of hourly global radiation sequences". Solar Energy, 23, pp. 47-51.

[36] B. Nguyen, and T. Pryor, "Generating artificial weather date sequences for Australian conditions". The 34th Annual Conference of ANZSES – Solar'96: Energy for Life, Darwin, Australia, pp. 101-108

[37] SOLCAST. Free Access to Solar Irradiance Data for Non-Commercial Research Use. https://articles.solcast.com.au/en/articles/2404457-free-access-to-solar-irradiance-data-for-non-commercial-research-use.

[38] HOMER Energy LLC. Hybrid Optimization of Multiple Energy Resources, Ver 3.14.0. HOMER Energy, Boulder, CO, USA, 2020.