Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2021.27117

222 (2021) 47–59
May

Model to generate solar radiation values for solar distillation numerical simulations

Bao The Nguyen[a,b,*], Viet Van Hoang[a,b,c], Hiep Chi Le[a,b]

[a]*Ho Chi Minh City University of Technology, 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh, Vietnam, Tel. (+84) 0906 33 11; emails: thebao@hcmut.edu.vn (B.T. Nguyen), hoangvanviethd@gmail.com (V. Van Hoang), lechihiep@hcmut.edu.vn (H.C. Le)*
[b]*Vietnam National University, 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh, Vietnam*
[c]*Ly Tu Trong College, 390 Hoang Van Thu Street, Tan Binh district, Ho Chi Minh, Vietnam*

Received 24 September 2020; Accepted 4 February 2021

ABSTRACT

Using solar stills to produce clean water for domestic use in developing countries is one of the affordable solutions. To avoid costly and time-consuming processes of solar still fabrication, numerical simulations are the perfect alternatives to predict the performance of the stills. The purpose of this study is to work out a model to generate the sequences of daily and hourly solar radiation in tropical countries that are very important to calculate or simulate any solar distillation systems. In this study, the modified 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 Graham is proposed to generate hourly clearness index sequences from generate daily clearness index series for these two locations. As being proved by some statistic configurations and the predicted distillate productivities of solar still simulations, two modified models are more accurate in predicting daily and hourly irradiances in comparison with original Aguiar's and Graham's model, respectively. Especially, the model to generate the sequences of hourly solar radiation 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 solar distillation simulation programs with very limited input parameters, including the latitude and monthly average daily clearness index values of the studied locations.

Keywords: Solar stills, Solar distillation numerical simulations; Monthly average daily solar radiation; Hourly solar radiation values; Daily solar radiation values; Markov transition matrix; MTM library

1. Introduction

Computer simulation programs can be very useful tools in the process of designing solar distillation systems. They enable users to optimize the performance of the solar stills 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 distillation 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

* Corresponding author.

1944-3994/1944-3986 © 2021 Desalination Publications. All rights reserved.

estimates of representative behaviour of the weather for the location under investigation. According to Duffie and Beckman [1], weather records of at least 8 y 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 y 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 for Europe [2]. 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 3-h 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 analyse 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 Topič [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 density function of the generated data are normal when stochastic models are used [10]. Gafurov et al. [11] incorporated spatial correlation of solar radiation 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 to model values of total solar radiation on horizontal surfaces, such as [12–16]. Ji and Chee [17] used a novel hybrid model of autoregressive and moving average (ARMA) and time delay neural network (TDNN) to predict hourly solar radiation in Singapore. The problem of these models is that they are "black boxes", and only mean values of daily global radiation have been analysed, leading to no significant information can be obtained [10]. Mora-Lopez et al. [18] 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.

Based on the above information, the stochastic approaches are still likely to be globally applicable. Therefore, in this study, the stochastic technique is chosen to generate daily and hourly global solar radiation from monthly average daily irradiances for solar distillation numerical simulations. First, a stochastic model is used to generate synthetic of daily irradiances from monthly average daily solar radiation values. Then the generated daily irradiance sequences are used to generate hourly solar radiation series.

2. Stochastic model to generate daily irradiances

2.1. Model to generate daily solar radiation sequences

Aguiar et al. [19] analyzed the daily radiation data for 300 months from nine stations having various types of climates. 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, $\bar{K}_t$, 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 Markov transition matrix (MTM) technique, Aguiar and his colleagues used the set of 300-month data to obtain ten 10 × 10 MTMs (called the MTM library). These MTMs were classified by different ranges of $\bar{K}_t$ values: one class for $\bar{K}_t \leq 0.30$, eight classes in steps of 0.05 between 0.30 and 0.70, and the last classes for $\bar{K}_t > 0.70$. 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 [20], 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, Macau, 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

[19], this model is believed to be more universally applicable to any location. Secondly, Mustacchi et al. [21] 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 [22]. 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 Topič [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 1 y in Ho Chi Minh City, representing for tropical forest climate (A_w) and 1 y in Da Nang City, representative for tropical monsoon climate (A_m). A pyranometer was used to measure global solar flux every 5 min, 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 h, 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 of Ho Chi Minh and Da Nang.

The 365 values of daily clearness index, $K_{T\text{mea}}$, for each of the two above cities were calculated. The daily clearness index is defined as [1]:

$$K_{T\text{mea}} = \frac{H}{H_0} \tag{1}$$

with H is the measured daily global solar radiation on the horizontal surface for Ho Chi Minh and Da Nang; H_0 is the daily extraterrestrial radiation, given by:

$$H_0 = \frac{24}{\pi} G_{SC} 3,600 \left\{ \begin{aligned} &\left[1 + 0,033.\cos\left(\frac{360n}{365}\right)\right] \times \\ &\left[\cos\varphi\cos\delta\sin\omega_s + \frac{\pi}{180}\omega_s \sin\varphi\sin\delta\right] \end{aligned} \right\} \tag{2}$$

In Eq. (2), G_{SC}, n, ϕ, δ and ω_s are solar constant, day of the year, latitude of the location, declination angle and sunset hour angle, respectively, 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]:

$$\overline{K_T} = \frac{\overline{H}}{\overline{H_0}} \tag{3}$$

with $\overline{H}$ 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 Eq. (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.

2.3. Application of Aguiar's model

Based on the data in Table 2, Aguiar's model is used to generate 365 values of daily clearness index $K_{T\text{gen.}}$ 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

Table 1
Monthly average daily global solar radiation on the horizontal surface $\overline{H}$ (kWh/m²) of Ho Chi Minh and Da Nang

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
HCM	3.61	4.99	5.04	5.20	4.65	4.83	4.80	4.90	4.43	4.07	4.16	3.85
Da Nang	2.87	5.22	5.17	6.15	6.36	6.65	5.64	5.20	4.79	4.12	3.31	2.34

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	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

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:

$$F_i\left(K_T\right) = \int_0^{K_T} P\left(K_T'\right) dK_T' \quad (4)$$

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 $K_{T\text{gen.}}$ is generated.

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

Some statistic configurations, including mean, median, min., max., standard deviation, 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 and 4, respectively.

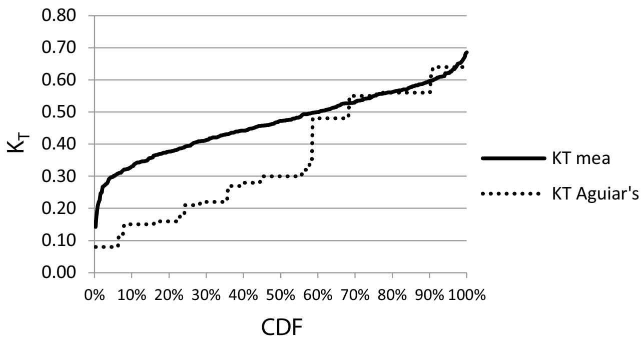

Fig. 1. Cumulative distribution function of K_T for Ho Chi Minh City.

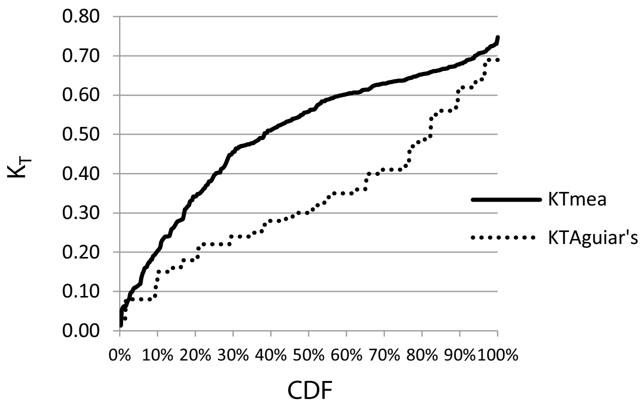

Fig. 2. Cumulative distribution function of K_T for Da Nang.

As shown in Figs. 1–4 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 31% and 57% for Ho Chi Minh City and 50% and 86% 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 [19]. Then, Aguiar and his colleagues suggested that the 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. 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 × 10 MTMs. In this study, besides the data from Aguiar's research, obtained from [23], 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

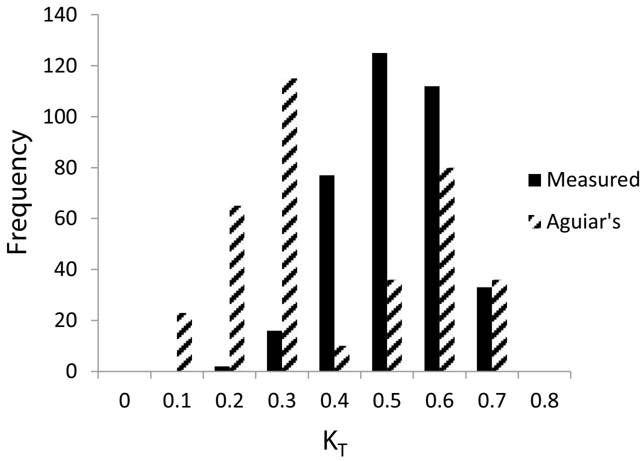

Fig. 3. Probability density function of K_T for Ho Chi Minh City.

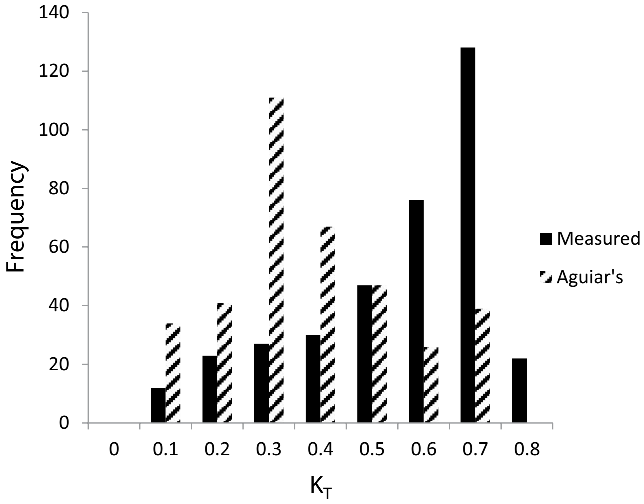

Fig. 4. Probability density function of K_T for Da Nang.

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

	Mean	Median	Min.	Max.	Std. dev.	MAE (%)	RMSE (%)
$K_{T\text{mea.}}$	0.47	0.47	0.14	0.69	0.10		
$K_{T\text{gen.}}$	0.36	0.30	0.08	0.64	0.18		
Error (%)	31	57				50.1	24.6

Table 4
Statistic configurations of K_T series of Da Nang

	Mean	Median	Min.	Max.	Std. dev.	MAE (%)	RMSE (%)
$K_{T\text{mea.}}$	0.50	0.56	0.01	0.75	0.18		
$K_{T\text{gen.}}$	0.33	0.30	0.05	0.80	0.17		
Error (%)	50	86				23.7	28

ranges of $\overline{K}_T$ values: one class for $\overline{K}_T \le 0.30$, 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. [19], but with smaller intervals between states to build ten non-square 20 × 10 MTMs

After the modification, the modified MTM library is used to regenerate the daily K_T values for Ho Chi Minh City and Da Nang. Figs. 5 and 6 graph the CDF of K_T generated by using the modified MTM library for Ho Chi Minh City and Da Nang respectively while Figs. 7 and 8 plot of the 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, MAE and 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 5 and 6, respectively.

Based on Figs. 5–8 and Tables 5 and 6, it can be concluded that the modified model has given much better results compared to those from Aguiar's model. As shown in Tables 5 and 6, the error percentages in mean and median of generated sequences are 1% and –4% for Ho Chi Minh City and 6% and 14% 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 30% and 53% for Ho Chi Minh City and 44% and 72% 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 because the original Aguiar model has been proved to be universally applicable [2,4,5,19,22]. As shown above, only the latitude and monthly average daily clearness index values of the studied locations are necessary for the generation.

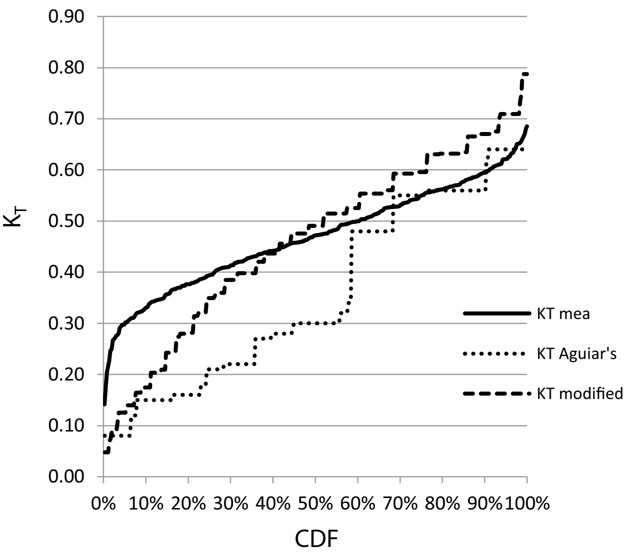

Fig. 5. Cumulative distribution function of three sets of K_T values for Ho Chi Minh City.

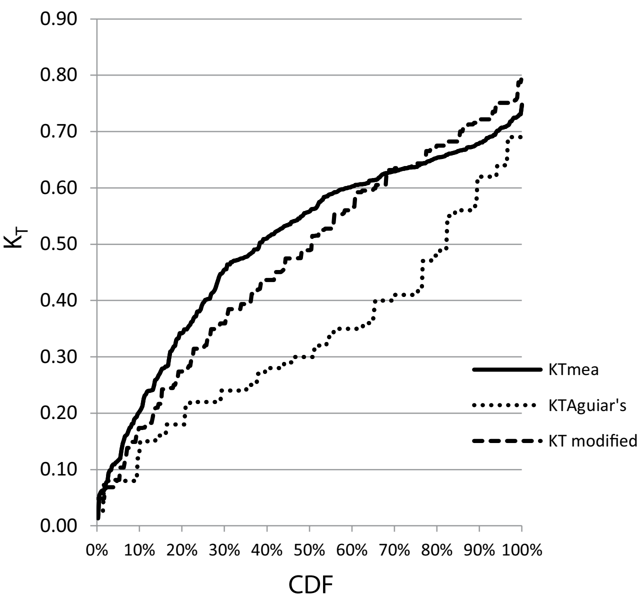

Fig. 6. Cumulative distribution function of three sets of K_T values for Da Nang.

3. Model to generate hourly irradiance sequences

3.1. Model to generate hourly solar radiation sequences

As mentioned above, there have been many researches using stochastic models to generate daily and hourly solar radiation sequences [2,4,5,22]. These researches were either based on Graham's approach [19] and Aguiar's approach [20].

For the stochastic model for the hourly radiation generation, Graham et al. [20] 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:

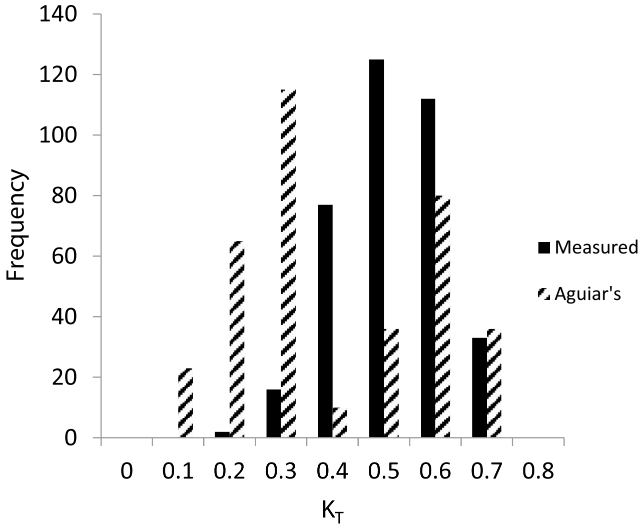

Fig. 7. Probability density function of three sets of K_T values for Ho Chi Minh City.

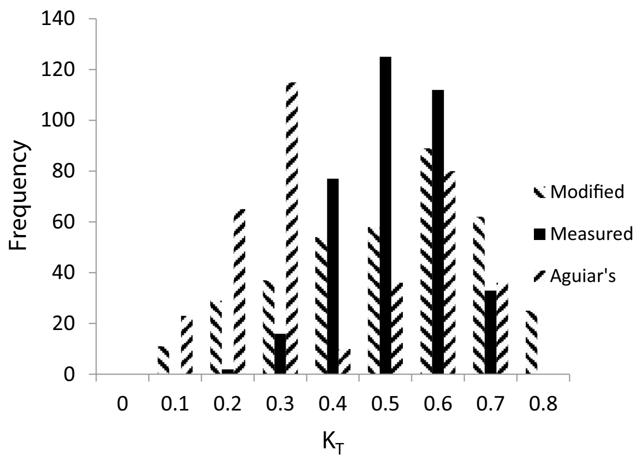

Fig. 8. Probability density function of three sets of K_T values for Da Nang.

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

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:

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

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

$$\kappa(K_T) = 1.141\frac{(1-K_T)}{K_T} \tag{7c}$$

Table 5
Statistic configurations of K_T values of Ho Chi Minh City

	Mean	Median	Min.	Max.	Std. dev.	MAE (%)	RMSE (%)
$K_{T\text{mea.}}$	0.47	0.47	0.14	0.69	0.10		
$K_{T\text{gen.}}$	0.46	0.49	0.05	0.78	0.18		
Error (%)	1	–4				16.8	8.1

Table 6
Statistic configurations of K_T values of Da Nang

	Mean	Median	Min.	Max.	Std. dev.	MAE (%)	RMSE (%)
$K_{T\text{mea.}}$	0.50	0.56	0.01	0.75	0.18		
$K_{T\text{gen.}}$	0.47	0.49	0.05	0.80	0.20		
Error (%)	6	14				4.5	5.1

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 [20]. Therefore, σ_α may be expressed as:

$$\sigma_\alpha(K_T) = 0.16\sin\left(\frac{\pi K_T}{0.90}\right) \tag{8}$$

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, ϕ = 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:

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

in which: Γ is the gramma function; u is the random variable normalized from hourly clearness index k_t 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 [20].

In developing another hourly radiation model, Aguiar et al. [19] 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:

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

with:

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

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

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 suggested a K_T dependent equation for ϕ as:

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

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.

Fernandez-Peruchena et al. [4] and Brecl and Topič [5] successfully used Aguiar's model to generate hourly solar irradiance in Spain and Slovenia, respectively whereas HOMER [24] uses Graham's approach to generate hourly solar radiation sequences in its software. Nguyen and Pryor [22] 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.

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

- The use of Gaussian mapping technique to treat the random component of k_t 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 Eq. (5) will be converted to non-Gaussian distribution η by the Norminv function in MATLAB. Then, the random components of k_t values are computed by:

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

with the standard deviation σ_a (K_T) calculated from Eq. (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 7 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.

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

Where Φ is the latitude of the location, L_{st} and L_{loc} 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; K_T [i][j] is the daily clearness index of ith day in the jth month; ω is the hour angle of the current hour; k_{tm} is the "long-term" average k_t values; σ_{kt} is the standard deviation of k_t about the long-term average k_{tm} value; ε_t is a random number from a Gaussian distribution; h is the hour considered; χ is a normally distributed stochastic variable with a mean of 0 and a variance of 1; 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.

Table 7
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%

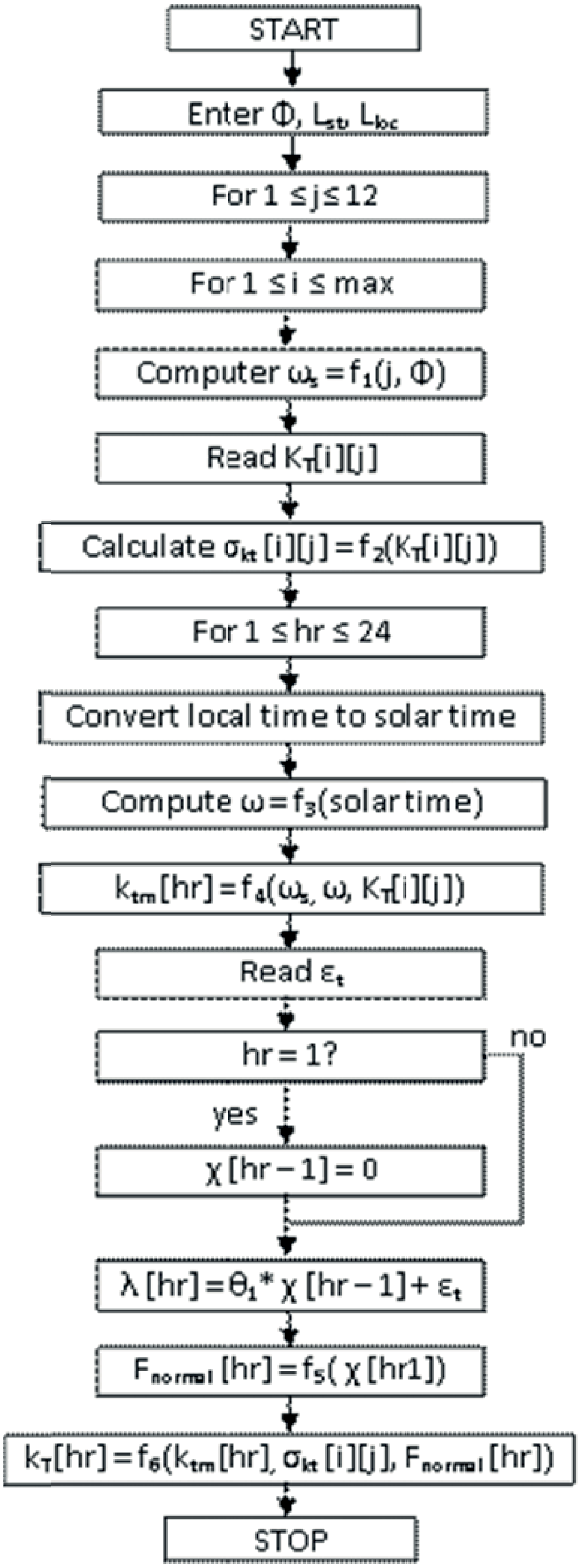

Fig. 9. The schematic diagram of the procedure to generate hourly k_t sequences.

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

3.2. Validating the generated hourly clearness index sequences

3.2.1. Comparing some statistic configurations

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:

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

with I is the measured hourly global solar radiation on the horizontal surface for Ho Chi Minh and Da Nang; I_0 is the hourly extraterrestrial radiation from hour angle ω_1 to ω_2, given by:

$$I_0 = \frac{12}{\pi} G_{\text{SC}} \times 3{,}600 \left\{ \begin{bmatrix} \left[1 + 0{,}033.\cos\left(\dfrac{360n}{365}\right)\right] \times \\ \left[\cos\phi\cos\delta\left(\sin\omega_2 - \sin\omega_1\right) + \right. \\ \left. \dfrac{\pi}{180}\left(\omega_2 - \omega_1\right)\sin\phi\sin\delta\right] \end{bmatrix} \right\} \tag{15}$$

Figs. 10 and 11 graph the CDF of hourly k_t for Ho Chi Minh and Da Nang respectively while Figs. 12 and 13 plot of the PDF of hourly k_t for these cities.

Some statistic configurations, including mean, median, min., max., standard deviation, MAE and RMSE of measured and generated k_t series of Ho Chi Minh City and Da Nang are also shown in Tables 8 and 9, respectively.

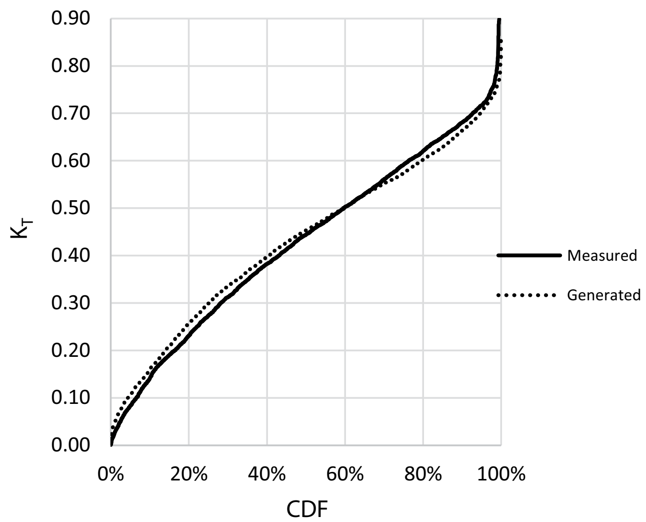

Fig. 10. Cumulative distribution function of hourly k_t for Ho Chi Minh City.

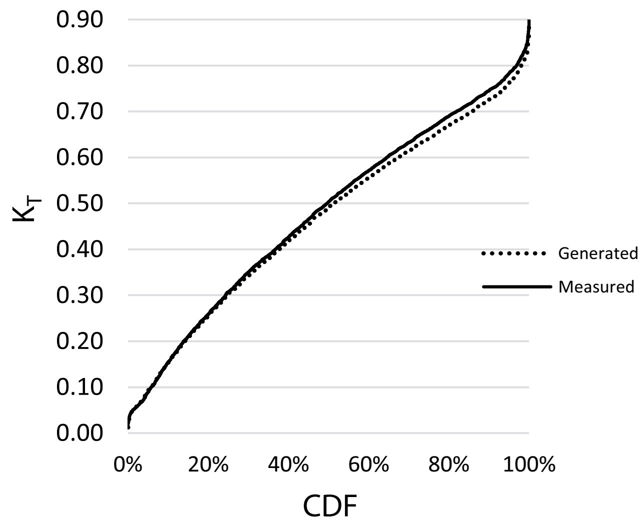

Fig. 11. Cumulative distribution function of hourly k_t for Da Nang.

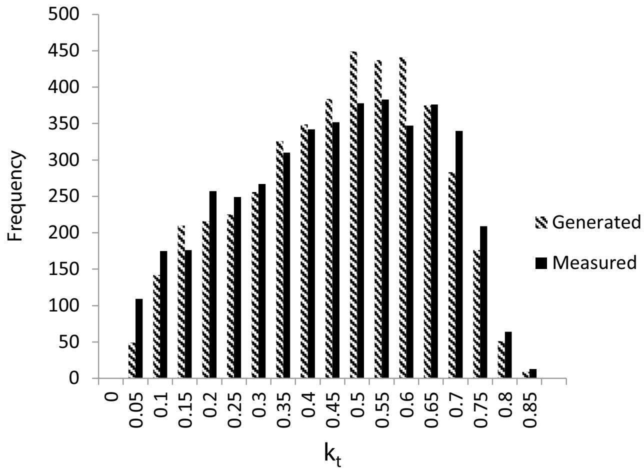

Fig. 12. Probability density function of hourly k_t for Ho Chi Minh City.

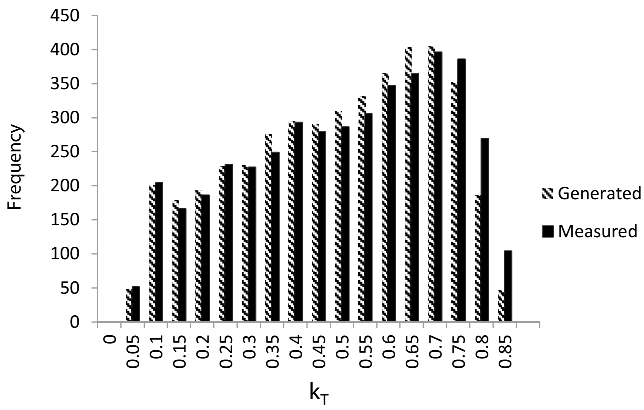

Fig. 13. Probability density function of hourly k_t for Da Nang.

Table 8
Statistic configurations of hourly measured and generated k_t series of Ho Chi Minh City

	Mean	Median	Min.	Max.	Std. dev.	MAE (%)	RMSE (%)
$k_{t\text{mea.}}$	0.426	0.443	0.001	0.899	0.197		
$k_{t\text{gen.}}$	0.433	0.453	0.007	0.858	0.191		
Error (%)	–1.5	–2.4				2.0	0.03

Table 9
Statistic configurations of hourly measured and generated k_t series of Da Nang

	Mean	Median	Min.	Max.	Std. dev.	MAE (%)	RMSE (%)
$k_{t\text{mea.}}$	0.459	0.491	0.019	0.905	0.210		
$k_{t\text{gen.}}$	0.465	0.492	0.012	0.884	0.217		
Error (%)	–1.3	–0.3				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 the world.

3.2.2. Using measured and generated solar radiation data to run solar distillation simulations

For further validating the solar radiation generating models, both measured and generated hourly irradiances of Ho Chi Minh City and Da Nang are used as input data to run SOLSTILL – simulation program for solar distillation systems [25]. SOLSTILL enables to simulate two types of solar stills: natural passive convection still (conventional solar still) and forced convection still with external condenser. Fig. 14 shows the heat and mass transfer processes in a conventional solar still. Fig. 15 presents a schematic diagram of a forced circulation solar still with enhanced water recovery and Fig. 16 shows the heat and mass transfer process in a forced circulation solar still [25].

The inputs for SOLSTILL include hourly solar irradiances, which are measured and generated in this study and hourly temperature averages for 12 months of a year that are achieved from National Center for Hydro-Meteorological Forecasting [26]. The outputs of SOLSTILL consist of hourly amounts of distillate water, hourly temperatures of the cover, basin water and the basin, etc. In this study, only hourly amounts of distillate water are considered. Then, daily amounts of distillate water are achieved by adding up the hourly ones for every day of the year. Finally, monthly average daily amounts of distillate water are calculated for comparison purposes in this study. Figs. 17 and 18 show the monthly average daily distillate productivity of a

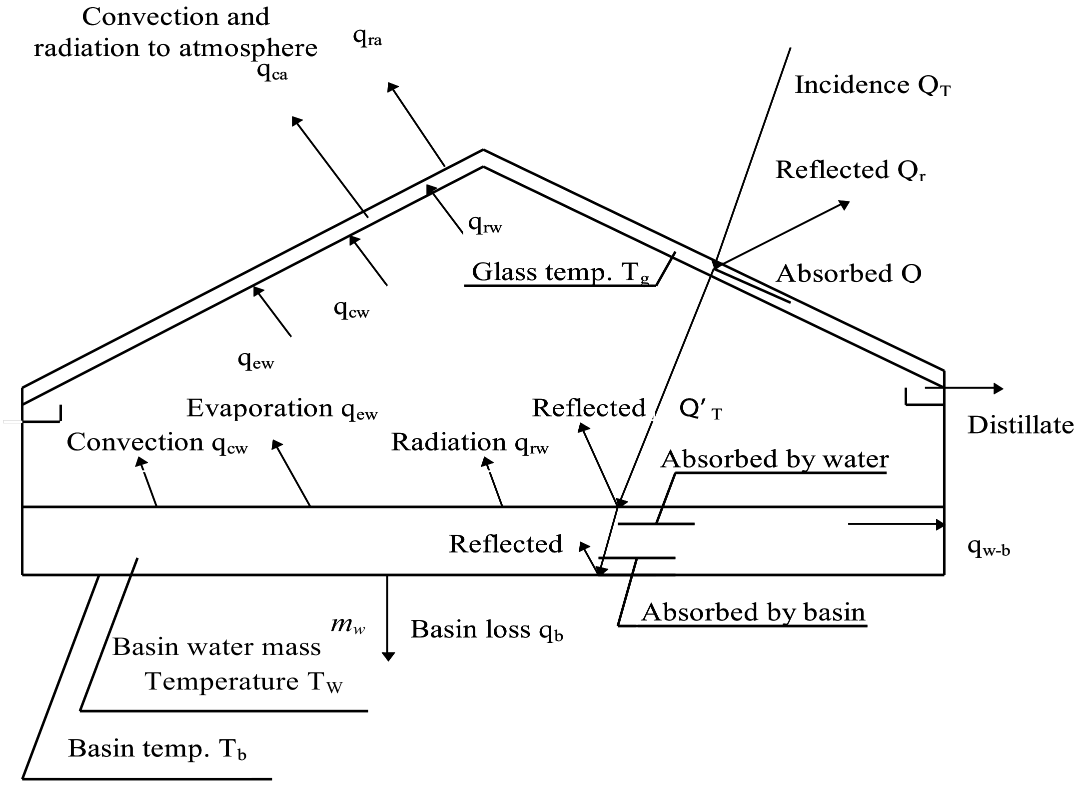

Fig. 14. The heat and mass transfer processes in a conventional solar still.

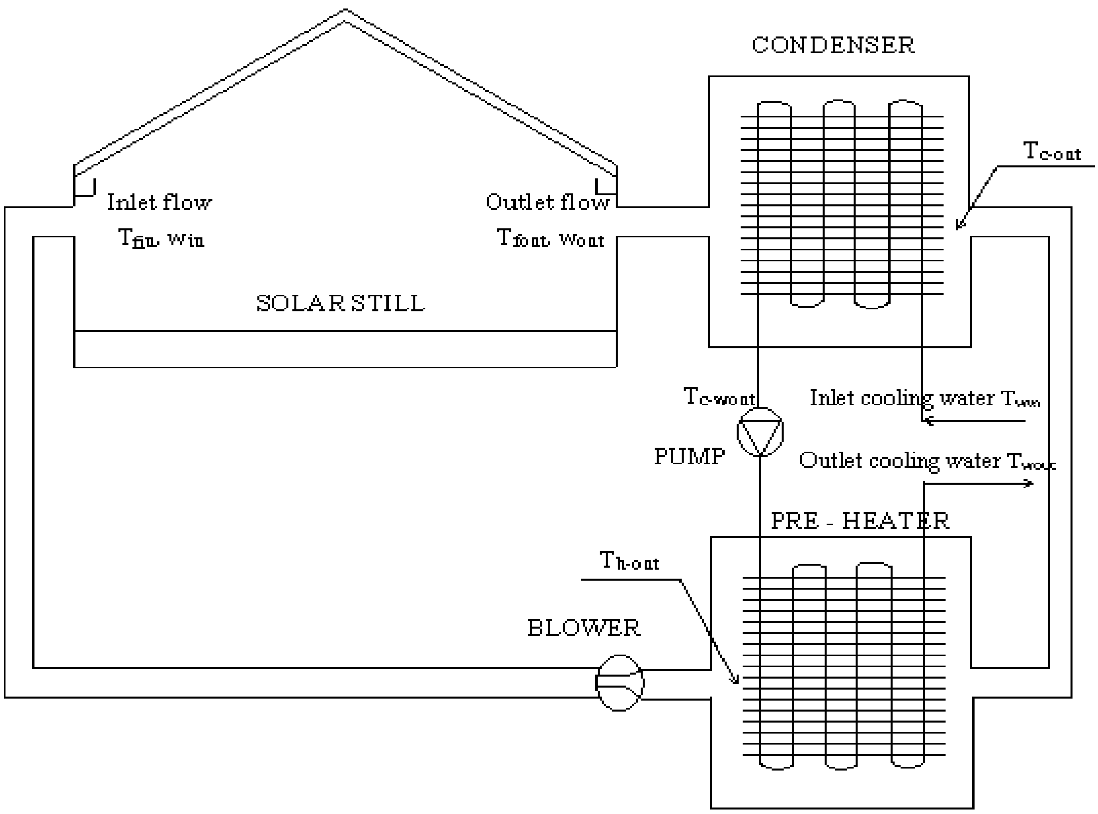

Fig. 15. Schematic diagram of a forced circulation solar still with enhanced water recovery.

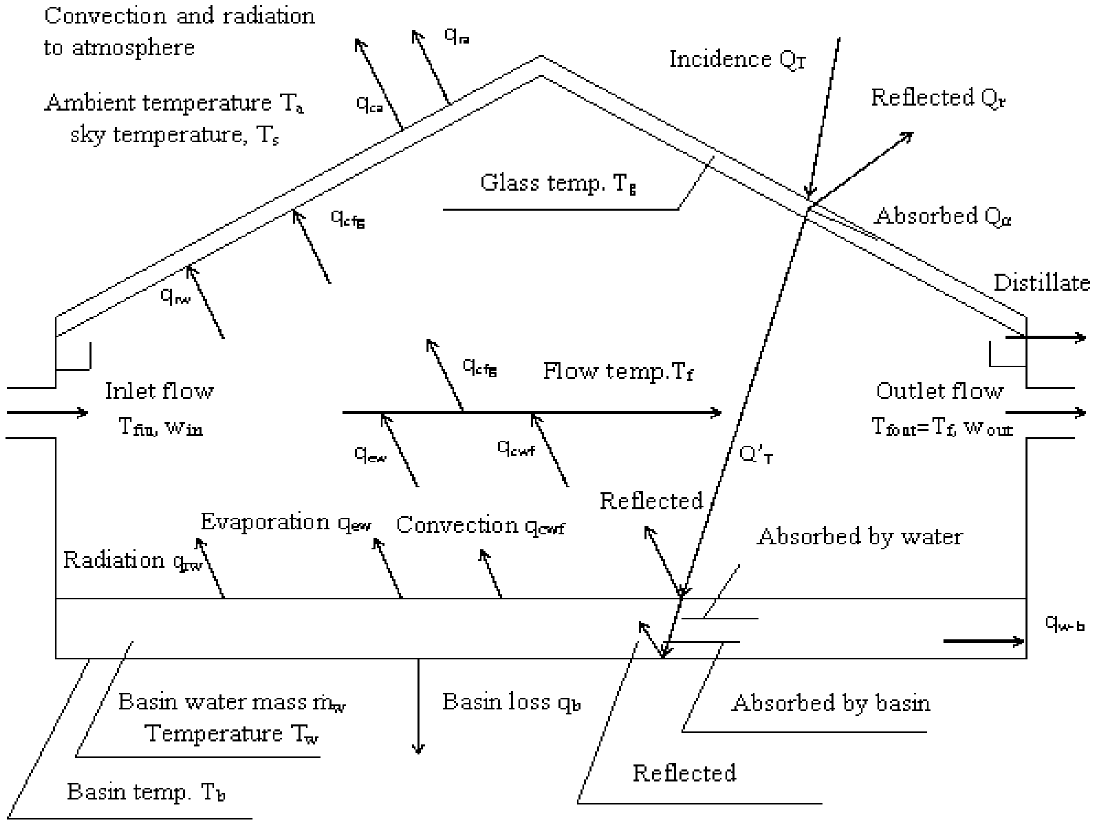

Fig. 16. The heat and mass transfer process in a forced circulation solar still.

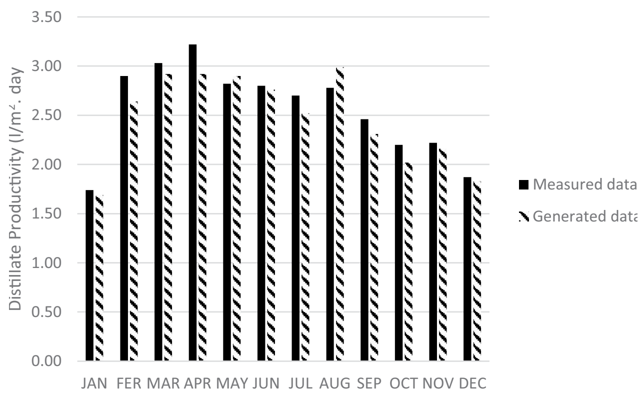

Fig. 17. Monthly average daily distillate productivity of a conventional solar still in Ho Chi Minh City.

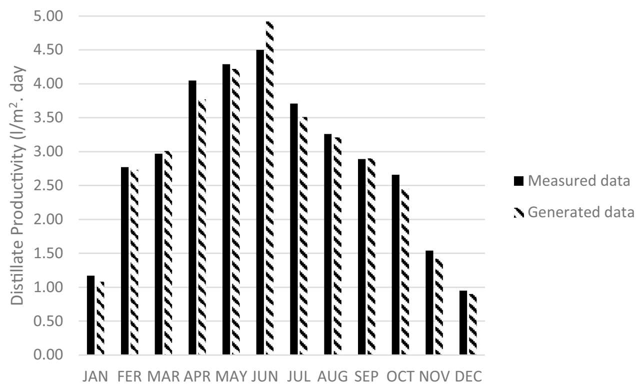

Fig. 18. Monthly average daily distillate productivity of a conventional solar still in Da Nang.

conventional solar still for Ho Chi Minh City and Da Nang whereas Figs. 19 and 20 show those of a forced circulation solar still with enhanced water recovery, respectively.

As shown in Figs. 17–20, the predicted monthly average daily distillate productivity of both a conventional passive solar still and a forced circulation solar still with measured irradiance data and generated solar radiation data as inputs are very closed. The largest error of predicted distillate with generated solar irradiances compared with those with measured data for both cities is 9.3%, occurred in April for a conventional solar still in Ho Chi Minh City. The errors of predicted yearly average daily distillate productivities between measured and generated solar irradiances for four above mentioned cases are less than 5%. Therefore, it can be concluded that the generated solar radiation data from the model of this study can be used to run any solar distillation simulation programs with acceptable accuracy.

4. Conclusion

In this study, the modified Aguiar's model was firstly used to generate daily clearness index series for Ho Chi

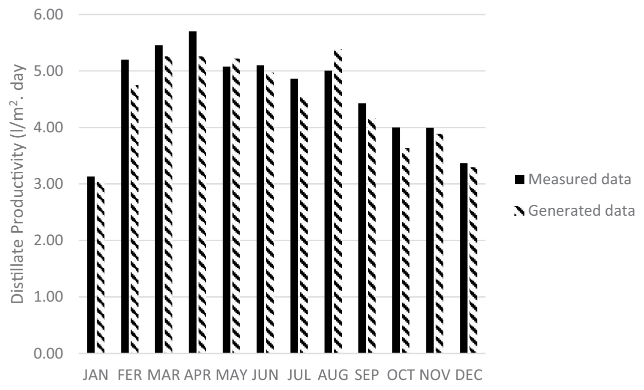

Fig. 19. Monthly average daily distillate productivity of a forced circulation solar still in Ho Chi Minh City.

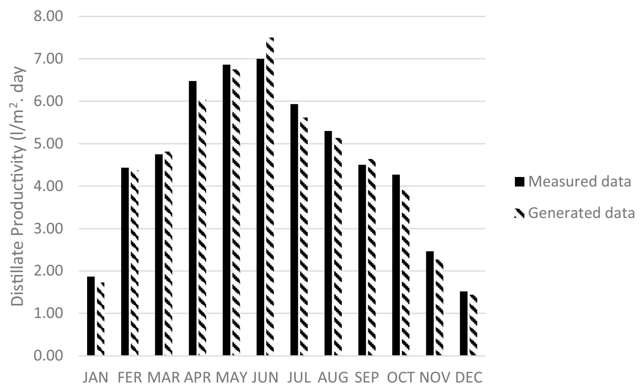

Fig. 20. Monthly average daily distillate productivity of a forced circulation solar still in Da Nang.

Minh City and Da Nang, two cities presenting for two climate types in tropical region. Then a modified model of Graham was proposed to generate hourly clearness index sequences from generate daily clearness index series for these two locations. Having been proved by some statistic configurations and the predicted distillate productivities of solar still simulations, two modified models in this study are more accurate in predicting daily and hourly irradiances in comparison with original Aguiar's and Graham's model, respectively. Especially, the model to generate the sequences of hourly solar radiation 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 solar distillation simulation programs with very limited input parameters, including the latitude and monthly average daily clearness index values of the studied locations.

Acknowledgement

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

References

[1] J. Duffie, W. Beckman, Solar Engineering for Thermal Processes, 4th ed., John Wiley & Sons, Inc., Hoboken, New Jersey, 2013.
[2] J. Boland, Time Series Modelling of Solar Radiation, V. Badescu, Ed., Modelling Solar Radiation at the Earth's Surface, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 283–312.
[3] B. Nguyen, 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, Ho Chi Minh City, Vietnam, 2017.
[4] C. Fernandez-Peruchena, L. Ramirez, I. Pagola, M. Gaston, Assessment of Models for Estimation of Hourly Irradiation Series from Monthly Mean Values, 15th SolarPACES Conference, Berlin, Germany, pp.121–126.
[5] K. Brecl, M. Topič, Development of a stochastic hourly solar irradiation model, Int. J. Photoenergy, 2014 (2014) 376504, doi: 10.1155/2014/376504.
[6] J.M. Bright, C.J. Smith, P.G. Taylor, R. Crook, Stochastic generation of synthetic minutely irradiance time series derived from mean hourly weather observation data, Sol. Energy, 115 (2015) 229–242.
[7] M. Hofmann, S. Riechelmann, C. Crisosto, R. Mubarak, G. Seckmeyer, Improved synthesis of global irradiance with one-minute resolution for PV system simulations, Int. J. Photoenergy, 2014 (2014) 808509, doi: 10.1155/2014/808509.
[8] T. Soubdhan, R. Emilion, Stochastic Differential Equation for Modeling Global Solar Radiation Sequences. Modelling, Identification, and Control, GOSIER, France, 2010, doi: 10.2316/P.2010.702-099.
[9] L. Magnano, J.W. Boland, R.J. Hyndman, Generation of synthetic sequences of half-hourly temperatures, Environmetrics, 19 (2008) 818–835.
[10] L. Mora-Lopez, A New Procedure to Generate Solar Radiation Time Series from Machine Learning Theory, V. Badescu, Ed., Modelling Solar Radiation at the Earth's Surface, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 313–326.
[11] T. Gafurov, J. Usaola, M. Prodanovic, Incorporating spatial correlation into stochastic generation of solar radiation data, Sol. Energy, 115 (2015) 74–84.
[12] W.I. Hammed, B.A. Sawadi, S.J. Al-Kamil, M.S. Al-Radhi, Y.I.A. Al-Yasir, A.L. Saleh, R.A. Abd-Alhameed, Prediction of solar irradiance based on artificial neural networks, Inventions, 4 (2019) 45, doi: 10.3390/inventions40300045.
[13] M. Vakili, S. Sabbagh-Yazdi, K. Kalhor, S. Khosrojerdi, Using artificial neural networks for prediction of global solar radiation in Tehran considering particulate matter air pollution, Energy Procedia, 74 (2015) 1205–1212.
[14] S.S. Priya, M.H. Iqbal, Solar radiation prediction using artificial neural networks, Int. J. Comput. Appl., 116 (2015) 28–31.
[15] F.S. Tymvios, S.Chr. Michaelides, C.S. Skouteli, Estimation of Surface Solar Radiation with Artificial Neural Networks, V. Badescu, Ed., Modelling Solar Radiation at the Earth's Surface, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 221–246.
[16] A. Sfetsos, A.H. Coonick, Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques, Sol. Energy, 68 (2000) 169–178.
[17] W. Ji, K.C. Chee, Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN, Sol. Energy, 85 (2011) 808–817.
[18] L. Mora-López, J. Mora, R. Morales-Bueno, M. Sidrach-de-Cardona, Modeling time series of climatic parameters with probabilistic finite automata, Environ. Modell. Software, 20 (2005) 753–760.
[19] R.J. Aguiar, M. Collares-Pereira, J.P. Conde, Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices, Sol. Energy, 40 (1988) 269–279.
[20] V.A. Graham, K.G.T. Hollands, A method to generate synthetic hourly solar radiation globally, Sol. Energy, 44 (1990) 333–341.
[21] C. Mustacchi, V. Cena, M. Rocchi, M. Stochastic simulation of hourly global radiation sequences, Sol. Energy, 23 (1979) 47–51.
[22] B. Nguyen, 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.

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

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

[25] B. Nguyen, The Mathematical Model of Basin-Type Solar Distillation Systems, Distillation - Modelling, Simulation and Optimization, Vilmar Steffen, IntechOpen, 2019, doi: 10.5772/intechopen.83228. Available at: https://www.intechopen.com/books/distillation-modelling-simulation-and-optimization/the-mathematical-model-of-basin-type-solar-distillation-systems

[26] National Center for Hydro-Meteorogical Forecasting. Available at: http://www.nchmf.gov.vn