The Libyan Journal of Agriculture
Volume (21), No. (1-2): 2016: 92-111

Original Article

Estimation of Sunshine Duration using Statistical Approach: Libya As A Case Study

Ahmed Ibrahim Ekhmaj and Milad Omran Alwershefani

Soil and Water Department, Faculty of Agriculture, University of Tripoli.

Abstract

Sunshine duration (SD) is an essential atmospheric indicator which is used in many agriculture, architects and solar energy applications. In many situations where data of sunshine duration may not be available due to temporal and financial constraints, developing alternative indirect methods based on theoretical considerations for determining SD are essentially required. In this study, seven models were developed using stepwise regression technique to estimate monthly sunshine duration for Libya. The predictors which were used as inputs differ from one model to another and they included monthly cloudiness index, total day length, mean relative humidity, depth of precipitation, mean maximum temperature, altitude and longitude over 16 meteorological stations spread across Libya during the period of 1961 – 2000 . The evaluation of the developed models was performed using a set of data of four meteorological stations representing different physiogeographic and climatic zones during 2001 and against Abdelwahed and Snyder (2015) equations which were developed for estimating sunshine duration for Libya. The statistical parameters of evaluation criteria included mean absolute error (MAE), root mean square error (RMSE), (RMSE %) and Nash and Sutcliffe Efficiency (NSE). The linear regression equation relating predicted with measured data with intercept equals zero and determination coefficient (R^2) were also used for evaluation purpose. According to the performance indicators, it was detected that six of the developed models were superior to the model with one parameter (cloudiness index) in estimating the sunshine hours. It was also found that all the developed models have better performance in estimating the sunshine duration as compared with Abdelwahed and Snyder (2015) equations. However, due to its few required variables, a model with two parameters (cloudiness index and total day length) is sufficient and can be used with confidence in estimating sunshine duration for Libya.

Keywords: Sunshine duration, Stepwise regression, Statistical model.

***Corresponding Author:** Ahmed I. Ekhmaj. Soil and water Dep., Faculty of Agriculture, University of Tripoli, Tripoli, Libya.
Phone. +218926577440. Email: *khmaj1@yahoo.com*

Received: 20/9/2016 Accepted: 01/12/2016

Introduction

Sunshine duration (SD) is one of the most important component of the atmospheric and meteorological components. Sunshine duration (SD) is defined by the World Metrological Organization (WMO) as the time during which the direct solar radiation exceeds the level of 120 W/m^2, and is normally measured in hours (Abdul-Alzzawi, 2006). This definition clearly reveals that SD can be used to determine the global solar radiation which in turn is valuable for agriculture, architects and solar energy applications (Velds, 1992). A lot of literatures assured its importance in various aspects of physical, chemical and biological processes. It suggested that SD can be used to characterize the climate of sites, especially health resorts and tourist destinations (WMO, 2008). Accurate solar radiation resource data are necessary at different steps of the design, simulation, and performance evaluation of any project involving solar energy (Gueymard and Myers, 2008). In irrigation practices, sunshine duration data are essentially used in estimating evapotranspiration. Sunshine duration can be also defined as the time during which the sun is visible. It is usually given in hours per day, month, season or year. Accordingly, the Campbell-Stokes instrument is used widely to determine the sunshine duration (Hinssen, 2006). Despite its pioneering in the task of determining the sunshine duration for longtime, its inaccuracies under some conditions have led to develop many alternative instruments. Such instruments include pyranometers and pyrheliometer. Pyranometers measure the global and diffuse irradiance while pyrheliometer measures the direct normal solar irradiance integrated over the entire solar spectrum. In many places where the instruments of measuring sunshine duration may not be available due to temporal and financial constraints, developing indirect methods based on theoretical considerations for determining SD are essentially required. There have been some studies conducted to estimate the SD from other weather data using regression analysis. Matzarakis and Katsoulis (2006) estimated mean annual and seasonal duration of bright sunshine over the Greek region from an empirical formula. The input parameters which they used in that formula depends on the distance of each station from the nearest cost, height above sea level, percentage of land cover around each station, latitude and longitude of each station. Yin (1999) presented a generic algorithm that captures global variation in monthly bright sunshine duration data in relation to temperature, precipitation and geographic location. The suggested method tends to be least robust for tropical and subtropical monsoonal Asia. Over the conterminous United States, the absolute error was within ±1.5 hr for 70% to 93% of the 60 monthly mean bright sunshine duration values for each of the stations. Umoh *et al.,* (2013) developed multiple linear regression models to estimate the monthly daily sunshine hours using four parameters i.e., Relative Humidity, Maximum and Minimum temperatures, Rainfall and Wind Speed, during a period of eleven years (1997 – 2007) for Calabar, Nigeria. Reddy (1974) has developed an empirical equation for estimating sunshine from total cloud amount. The results obtained with this equation were compared with observed values and were found to be in good agreement. Robaa (2008) developed a simple regression formula to estimate relative sunshine duration, using readily available observed data of cloud amount, in Egypt. He concluded that formula gives precise estimations for relative sunshine duration and it was recommended for use at any location in Egypt. In north China, Yang *et al.,* (2009) found that the monthly sunshine had a high correlation with precipitation, relative humidity and wind speed. Recently, Abdelwahed and Snyder (2015), derived three simple regression equations based on monthly averages of daily mean temperature, relative humidity, and wind speed, from 16 meteorological stations to estimate the sunshine hours in Libya.

Although these equations were successfully used to predict the evapotranspiration in California-USA, a question can be raised about the ability of these simple equations which were developed based on one climate variable to describe the sunshine duration in area with about 1.76 million square kilometers such as Libya which is different in terrain and climatic conditions. There are several important climatic and spatial variables that have been successfully used in the estimation of the SD which may lead to improve the estimations if it has been taken into account. This study aimed to evaluate the applicability of MLR to estimate the SD under the Libyan conditions.

Materials and Methods

The study area and dataset

Libya is situated in North Africa between longitude of 9° 25` E and 25° E and between latitude of 19 °27` N and 33° 10` N. It covers a total area of 1.76 Million km^2. Desert which covers more than 95% of the country represents a dominant feature of Libya. Libya climate is determined by contrasting Mediterranean and Sahara climates (El-Tantawi, 2005).

According to Köppen's climatic classification, Libya has three climatic zones i.e., a hot desert climate which covers a large percentage of the mid- and southern parts, a hot steppe climate type in northern parts and Mediterranean zone in northeastern Libya (MÜHR, 2004). Such variability in climatic zones is attributed to the interactions of the location and physiographic features of Libya. Therefore, it was assumed as a matter of fact, that such variations in the climate, location and physiographic features would affect the sunshine duration over all the area of Libya. Sixteen meteorological stations which comprised all types of climatic zones throughout Libyan territory were chosen to achieve the objectives of this study. Figure 1 shows the map of Libya, indicating the stations which were used for the study.

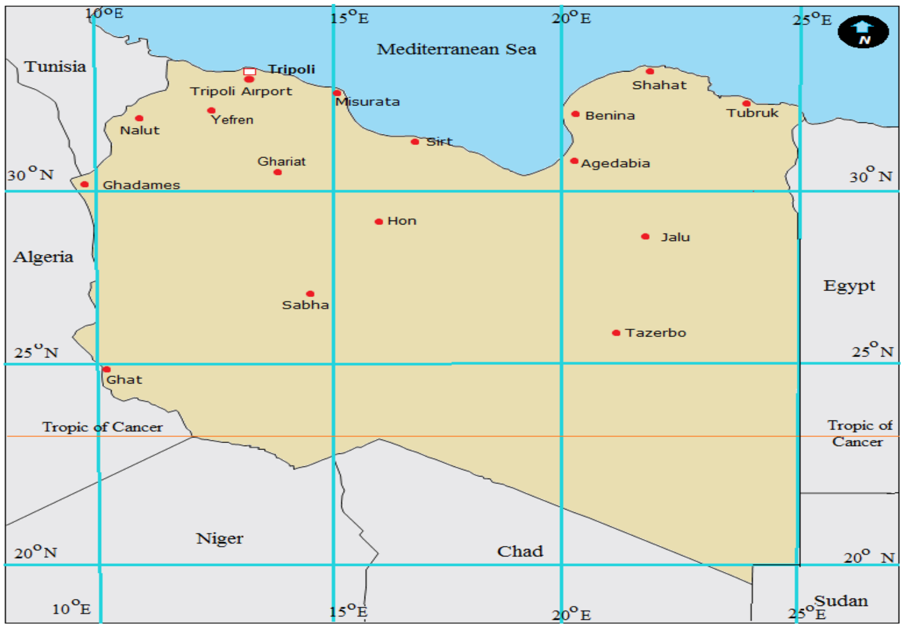


Figure 1. Location of Libya and representative meteorological stations.

The measured monthly mean data of sunshine duration hours (h), maximum temperature (max T, C^{o}), minimum temperature (min T, C^{o}), mean temperature (mean T, C^{o}), relative humidity (RH, %), total day length (N, hr), precipitation depth (P, mm) and cloudiness index (CI, octas) for the stations were collected for 30 years, from 1961 until 2001. The duration of sunshine hours were measured by the mean of Campbell-Stokes type sunshine recorder. The geographical coordinate's data as represented by latitude (LAT) and longitude (LONG) and the elevation above the sea level (Elev, m) of each station were also collected and considered for estimating sunshine duration (Table 1). These data were obtained from the archives of the Libyan National Meteorological Center.

Multiple regression analysis approach was implemented for estimating sunshine duration. Modeling of multiple linear regressions assumes a linear relationship (direct line) between the dependent and the predictor variables. Two main conditions were suggested in order to perform the multiple regression analysis. The first was to build several multiple regression equations based on the available data of climatic variables whereas the related incomplete and inaccuracy data were excluded. This suggestion was made to overcome the lack of climatic data spatially and temporally (missed or not measured). The second condition was associated with obtaining general multiple regression equations taking into consideration the latitude (Lat) and the longitude (Long) coordinates of meteorological stations wherever it was possible (i.e. statistically).

The suggested multiple linear regression equation can be expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \varepsilon \qquad (1)$$

Where Y is the dependent variable, βo is the intercept, β_1, $\beta_2 .....\beta_n$ are the regression

Table 1. Geographical coordinates and elevation of the stations.

Station	WMO No.	Longitude (E)	Latitude(N)	Elevation (m)
Ghadames	62103	09.50	30.13	346
Ghat	62212	10.15	25.13	692
Nalut	62002	10.98	31.86	621
Yefren	62008	12.55	32.08	691
Tripoli Airport	62010	13.15	32.67	081
Ghariat	62120	13.58	30.38	497
Sabha	62124	14.43	27.02	432
Misurata	62016	15.05	32.32	032
Hon	62131	15.95	29.12	263
Sirt	62019	16.58	31.20	013
Agedabia	62055	20.17	30.72	007
Benina	62053	20.27	32.08	130
Tazerbo	62259	21.13	25.80	261
Jalo	62161	21.57	29.03	045
Shahat	62056	21.85	32.82	649
Tubruk	62062	23.92	32.10	050

coefficients, and X_1, X_2,.... X_n are the independent variables and ε is the error term which was distributed normally with zero mean and variance S^2. Accordingly, the mean monthly values of sunshine hours represented the dependent variables while the mean monthly values of the longitudinal (decimal degrees), latitude (decimal degrees), monthly mean min T (°C), max T (°C), mean T (°C) , mean RH (%), monthly precipitation (P, mm), monthly cloudiness index (CI, octas), monthly mean day time N (hrs) and Elevation (m) for sixteen meteorological stations were considered to be as independent variables. The parameters of β_0, $\beta_{1..}$, B_n were estimated by means of Stepwise regression technique using the sample sets for the period from 1961 to 2000. Statistically, two assumptions about the populations have to be fulfilled when applying regression. The dependent variables have to be normally distributed and the observations of the independent variable should be statistically independent (Tveito, 1998). A histogram of the residuals was adopted to check the assumption of normality of the error term. The shape of the histogram should approximately follow the shape of the normal curve. In addition, the diagnostic checking in the regression relationships excludes and removes the highly correlated variable with other X variables from the multiple regression equation. Such procedures ensure the acceptance of the regressions models. In this regard, stepwise regression technique, which was adopted in this study to develop the models, removes and adds variables to the regression model for the purpose of identifying a useful subset of the predictors. All data were processed and loaded into the statistical Software SPSS®, Version 20, (SPSS, 2011).

Verification and Evaluation Criteria:

Monthly meteorological data of four different stations located in various zones during 2001 were selected to verify and assess the performance of the developed multiple regression equations. These stations included Tripoli Airport, Sirt, Tubruk and Sabha. They created advantages which could help for verification purpose since they were widely distributed throughout the Libyan territory. In addition, several statistical parameters were used to evaluate the accuracy of the predicted models. These parameters included mean absolute error (MAE), root mean square error (RMSE), percentage root mean square error ((RMSE %) and Nash and Sutcliffe Efficiency (NSE). These parameters were calculated as stated by (Nash and Sutcliffe, 1970, Ekhmaj *et al.*, 2013 and Kandirmaz, 2014) as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|(M_i - P_i)| \qquad (2)$$

$$RMSE = \left(\sum_{i=1}^{n}(M_i - P_i)^2/n\right)^{1/2} \qquad (3)$$

$$RMSE\ \% = \frac{RMSE}{\bar{\bar{M}}} \times 100 \qquad (4)$$

$$NSE = \left(1 - \frac{\sum_{i=1}^{n}(M_i - P_i)^2}{\sum_{i=1}^{n}(M_i - \bar{\bar{M}})^2}\right) \times 100 \quad (5)$$

where, M_i is the measured sunshine hours, P_i is the predicted sunshine hours, n is the number of records and $\bar{\bar{M}}$ is the mean value of the measured sunshine hours. The MAE, *RMSE* and *RMSE %* statistics have as lower limit the value of zero, which is the optimum value for them, while the optimum value of NSE is 100%.

The linear regression equation with intercept equals zero in combining determination coefficient (R^2) were also used to determine the goodness of the predicted model. The general form of this equation can be expressed as:

$$y = \alpha\, x \qquad (6)$$

Where y represents the predicted data, α represents the slope of the regression line and x represents the measured data. For perfect fit the value of the slope equals one (Ekhmaj *et al.*, 2013). However, the MAE, RMSE and RMSE %, NSE, α and R^2 were also adopted to diagnose the developed regression equations during construction stage of modeling.

To test the developed models more thoroughly, a sole published study of determining sunshine duration in Libya as conducted by Abdelwahed and Snyder (2015) were suggested for comparison purpose. The models were derived based on linear regression analysis, in which the sunshine duration could be described as a function of monthly averages of air temperature (T, °C),mean relative humidity (RH, %), wind speed at elevation of 2 meter above the ground (U_2, m s^{-1}) as follows:

$$SD = 4.352 - 0.232T \qquad (7)$$

$$SD = 24.035 - 0.283RH \qquad (8)$$

$$SD = -0.715 + 2.501U_2 \qquad (9)$$

Results and Discussion

Input data characteristics:

The descriptive statistical characteristics of annual mean maximum temperature (max T), mean minimum temperature (min T), mean temperature (mean T), cloudiness index (CI), precipitation, sunshine duration (SD) and total day length (N) were determined using the data measured at mentioned weather stations from 1962 to 2000 (Table.2 and 3). Notable big values of coefficient of variation (CV) and standard error (SE) in the data of annual climatic variables can be attributed to their spatial and seasonal variations. Annual maximum temperature (max T) ranges from 9 °C at Shahat in the north east during July to 45 °C at Ghdamas in the south west during January. Such variations were also detected in the annual (min T). It ranges from -0.5 °C to 29.4 °C at Ghat during January and August, respectively. It can be also observed that the mean annual temperature (mean T) ranges from 6.2 °C at Shahat during January to 35.7 °C at Ghat during June. El-Tantawi (2005) indicated that Mediterranean and Sahara climates contrasting Libya's climate whereas the air masses of either continental or maritime origin affect climate.

from 6.2 °C at Shahat during January to 35.7 °C at Ghat during June. El-Tantawi (2005) indicated that Mediterranean and Sahara climates contrasting Libya's climate whereas the air masses of either continental or maritime origin affect climate.

Annual relative humidity (RH) averaged 52.5% and ranged from 13% at Ghat and Ghdamas in the south during June and July to 97% at Tubruk in the east coast of Libya during August. The variations in the annual relative humidity are owing to the influences of the Mediterranean Sea in the coastal area as compared with area which dominated by desert climate type.

Mean annual precipitation (P) over all Libya averaged 13 mm. Most of precipitation occurred during winter and varied dramatically from north to south. It ranged from 0 mm at Ghat during August to 434 mm at Shahat during January. Unlike other climate variables; the mean value of annual precipitation is obviously greater than its median. Such behavior is normally associated with the precipitation and indicating that the skewness will be positive and have long right tails with almost zero (Sen and Eljadid, 1999).

The mean annual cloudiness index (octas) decreases southwards. It ranges between 0.0 at Ghadames in July and 6.6 at Shahat in January.

The increased cloudiness of the Mediterranean region occurs during winter when the climate is dominated by cyclonic storms conditions. The invasion of extratropical systems in winter from the north passing over the Mediterranean, which

Table (2) Descriptive statistical characteristics of meteorological data

Variable	Mean	Max	Min	Median	ST DE	C.V (%)
Max T (^{o}C)	27.0	45.3	9.0	27.1	7.2	26.8
Min T (^{o}C)	14.5	29.4	-0.5	14.8	6.0	41.7
Mean T (^{o}C)	20.7	35.7	6.2	21.0	6.5	31.1
Mean RH (%)	52.5	97.0	13.0	54.0	16.6	31.6
P (mm)	13.0	434	0.0	0.0	29	228
CI (octas)	2.1	6.6	0.0	2.0	1.3	61.5
N (h).	12.0	14.3	10.5	12.1	1.3	11.2
SD (h)	9.1	13.5	3.0	9.0	2.0	22.0

Table (3) Characteristics of most extreme values for each meteorological variable

Variable	Month with		Station with	
	maximum value	minimum value	maximum value	minimum value
max T (^{o}C)	July	January	Ghadames	Shahat
Min T (^{o}C)	August	January	Ghat	Ghat
Mean T (oC)	June	Januray	Ghat	Shahat
Mean RH (%)		June		
	August	July	Tubruk	Ghat Ghadames
P monthly(mm)	December	June	Shahat	Ghat
CI (octas)	January	July	Shahat	Ghadames
N (h).	June	January	Shahat	Shahat
SD (h)	July	December	Jalo	Shahat

causes increased cloudiness were also noticed in similar sites (in northern Egypt) and gradually decreased toward the south (El-Metwally, 2005). However, Martyn(1992) indicated that the weather conditions in the Sahara, are not conductive to cloud formation.

The distribution of annual Maximum possible duration of sunshine hours or daylight hours (hour) was investigated. It ranges from 10.5 hour in Shahat during January to 14.3 hours in Shahat during June. The coefficient of variation 11.2% is owing to the variations in latitude of the locations and the sun declination .i.e., the day of the year (Duffie and Beckman, 1991).

Sunshine hours are site and time dependent (Karume *et al.*, 2007). Figure (2) shows the pattern distribution of the mean sunshine duration during March (a), June (b), September (c) and December (d). It gives a clear vision about the variations of sunshine duration pattern in time and space. The figures indicate that the intensity of sunshine duration values tend to increase southward and decrease northward of Libya during March and December. Such distribution pattern reveals that the sunshine duration depends on Latitude. On the other hand, the intensity of sunshine duration values tend to decrease westward and increase eastward of Libya during June and September and reveals that the sunshine duration depends on Longitude.

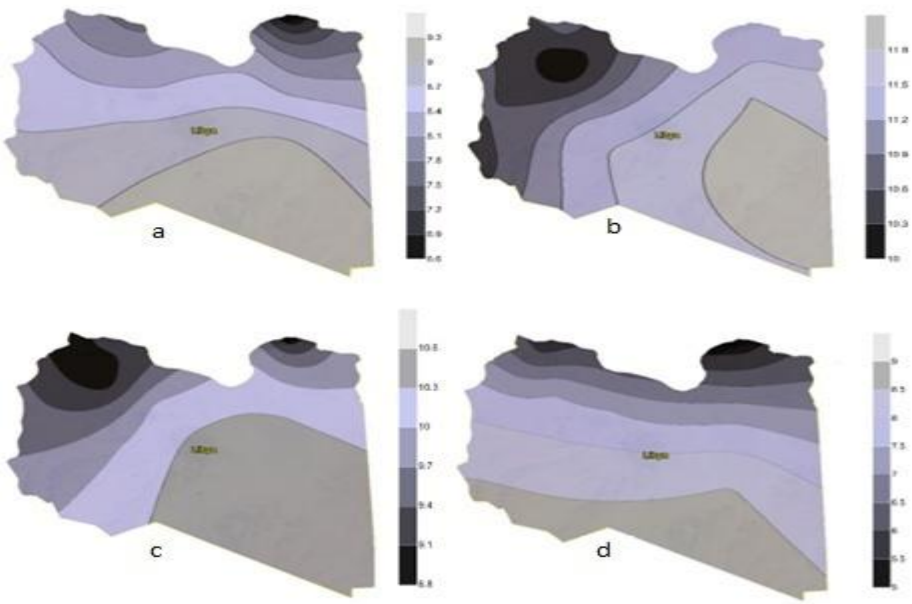


Figure. 2. Distribution of the measured sunshine duration (h) over Libya during the months of March (a), June (b), September (c) and December (d)

The mean annual sunshine duration averaged 9.1 hours with maximum value of 13.5 hours measured in Jalo during July. The minimum value 3.0 hours was recorded in Shahat during January. This low value is due mainly to the prevailing rainy/cloudy conditions during January. The occurrence of the maximum values of sunshine duration during summer could be almost attributed to the missing of depressions developing across Mediterranean region. In winter when the value of sunshine duration reaches its minimum, the cyclonic (depressions) and frontal activity show their maximum frequency (Matzarakis and Katsoulis, 2006).

Overall, as it was mentioned by (Martina et al., 2014), Libya's climate is mainly determined by contrasting Mediterranean and Sahara climates.

Sunshine Duration Modeling

A multiple regression molding was adopted in order to estimate the sunshine duration. Essentially, it was proposed to construct multiple regression equations by varying the type and the number of parameters data. Accordingly, Stepwise regression technique was implemented to determine the sunshine hours with less constraints of data.

In order to measure the degree of linear dependency between the sunshine duration and other variables, the correlation coefficient of Pearson correlation coefficient was determined. The correlation coefficient of Sunshine duration with the other variables were significant at the 0.05 level with values of -0.240, -0.014, 0.760, 0.682, 0.745, -0.431, -0.544, -0.764, -0.760, -0.02 for Lat, Long, max T, min T, mean T, Mean RH, P, CI, N and Elevation (Elev), respectively. The negative signs of the coefficients reveal that the values of sunshine duration increased with the decreasing of the values of Long, max T, Mean RH, CI, N and Elev.

The seven possible multiple regression equations were derived based on the available metrological data, elevation and coordinates of 16 stations scattered over the Libyan area (Table 4). As depicted by the shape of the normality of the error term for the developed models, it was found that the histograms approximately follow the shape of the normal curve and reveal that all residuals of the developed models are well- behaved, Appendix (1). It is pertinent for comparison purposes; the equations were numbered and sorted in ascending order according to the number of input parameters. The results of the regression analysis which were extracted from the analysis of variance (ANOVA) table and evaluation parameters for predicting sunshine hours are presented in Table (5).

It can be observed from the values of F ratio and p that all developed equations are of highly statistically significant effect at $\alpha = 0.05$ (F-test p-values < 0.05). The p-values in the ANOVA table (0.000) show that the models estimated by the regression procedure are significant at α -level of 0.05. The standard error of the estimate (SE) is a measure that shows how much the estimated average obtained is accurate (Khalifeloo *et al*., 2015). It can be noted from table (5) that the lower value of (SE) (0.806 h) was found by model 7. The higher (SE) value (1.287 h) was found by model 1.

The lower value of determination coefficient (R^2) was 0.58 as obtained from model 1 while the models 6 and 7 have the higher value of determination coefficient (0.84). The lower and the higher values of the determination coefficient (R^2) indicate that the predictors explain 58% and 84% of the variance in the sunshine duration, respectively. The results of ANOVA data reveal that the derived equations can be used to estimate the sunshine duration using the described predictors. In order to assess the performance of the developed equations α, MAE, RMSE, RMSE % and NSE were determined for each equation. As mentioned in table (5), it was found that the values of α, MAE (hr), RMSE (hr), RMSE % and NSE % ranged from 0.98 and 0.99, 0.60 h to 1.03 h, 0.81 h to 1.29 h, 8.87 % to 14.17% and 58.43% to 83.72%, respectively. According to the performance indicators, it can be detected that model 7 is superior to other models in estimating the sunshine hours. Notably, the results presented in table 4 and 5 indicated that the performance of the equations were enhanced by adding new appropriate variable. For instance, the performance of model 1 which is based only on cloudiness index enhanced more thoroughly by adding the N variable (i.e., model 2).

Table 4. Deduced sunshine duration models

Model	inputs	Expression
1	1	SD= 11.567- 1.161 Cl
2	2	SD=0.999 - 0.827Cl + 0.808 N
3	3	SD =0.466 - 0.892 CI + 0.821 N + 0.010 Mean RH
4	4	SD =0.543 - 0.849 CI + 0.807 N + 0.011 Mean RH - 0.004 P
5	5	SD =0.523 - 0.807 CI + 0.726 N + 0.014 Mean RH - 0.005 P + 0.027 Max T
6	6	SD =0.308 - 0.810 CI + 0.712 N + 0.017 Mean RH - 0.005 P + 0.033 Max T+0.0003 Elev
7	7	SD= 0.10 – 0.806 Cl + 0.703 N+ 0.016 Mean RH – 0.005 P + 0.034 Max T + 0.00045 Elev + 0.019 Long

Table (5) Analysis of variance (ANOVA) and evaluation parameters for deduced predicting sunshine duration models

Model	F- ratio	p	SE (hr)	R^2	α	MAE (hr)	RMSE (hr)	RMSE %	NSE
1	7916.3	0.00	1.287	0.58	0.98	1.03	1.29	14.17	58.43
2	13272.5	0.00	0.835	0.83	0.99	0.62	0.84	9.19	82.50
3	6210.4	0.00	0.825	0.83	0.99	0.61	0.83	9.08	82.93
4	6969.0	0.00	0.818	0.83	0.99	0.61	0.82	9.01	83.19
5	5678.9	0.00	0.812	0.83	0.99	0.61	0.81	8.95	83.43
6	4774.6	0.00	0.809	0.84	0.99	0.61	0.81	8.91	83.57
7	4132.3	0.00	0.806	0.84	0.99	0.60	0.81	8.87	83.72

Verification of the developed sunshine models:

To determine the reliability of the developed equations, data other than those used in the identification and developing steps are needed. The required inputs data of Tripoli airport, Sirt, Tubruk and Sabha during the year of 2001 were used to test the developed equations more thoroughly.

The comparisons between the measured monthly sunshine duration and those predicted from the developed models are shown in Figures 3, 4, 5 and 6. As it can be seen from the figures, there is a high accordance between measured and predicted data of monthly sunshine duration by all models except model 1 for Tripoli airport, Sirt, Tubruk, whereas the measured data were closely matched by the predicted. This is due to the fact that models 2 to 7 have additive variables.

In Sabha, a higher accordance between measured and predicted data by model 1 compared with other models occurred during the rainy season (Sep – April).

Such result reveals that the dominant factor controlling the monthly sunshine duration in the rainy season in Sabha is cloud index (octas). In the other months, the predictions are slightly enhanced by additional variables.

Using Abdelwahed and Snyder (2015) equations in predicting sunshine duration versus measured and predicted data which were obtained from the models 1 to 7 were performed and plotted in the figures (3), (4), (5) and(6).

It can be noted that the predicted data of sunshine duration as a function of wind speed in Sirt are included in figure (4) due to lack of wind speed data in Sirt during August, September and October in 2002. Consequently, applying equation (9) was not performed properly.

Higher discrepancies can be noted between the measured and the calculated values using equations (7), (8) and (9) as compared with the models 1 to 7. However, among Abdelwahed and Snyder (2015) equations, it is likely that equation (7) performs well in estimating sunshine duration in Tubruk and Sabha.

For further analysis, the performance indicators, i.e., α, R^2, MAE (hr), RMSE (hr), RMSE (%) and NSE (%) were also used to test how well the results predicted by the different models mimic the measured sunshine duration.

The results of this analysis are summarized in Table. 4. It can be seen that the slope values (α) of the best fitted line of the models 1 to 7 ranges from 0.948 to 0.999. The slope values reveal that the percent error between the measured and predicted data is in the range of -5.2 % to -0.1%.

In all models except model (1) in case of Sirt, the proportion of the total variance of the measured data explained by the predicted data as represented by coefficient of determination R^2 are in the range of 0.787 to 0.910.

The negative sign of determination coefficient as found in applying model (l) in Tripoli airport, Tubruk and Sabha indicates that the linear relationships between the measured and predicted data do not follow the linear trend as described by equation (6). The negative coefficient of determination is reported for a model which fits worse than models consisting only of the sample mean (Startz, 2013).

Cameron *et al*, 2011 indicate that R^2 can take negative values when the model does not contain a constant i.e., intercept term. However, examining equation 6 which used for evaluation purpose, leads to indicate that equation does not include the intercept term. As it is indicated previously, it was found that the models (2) to (7) performed well in predicting sunshine duration.

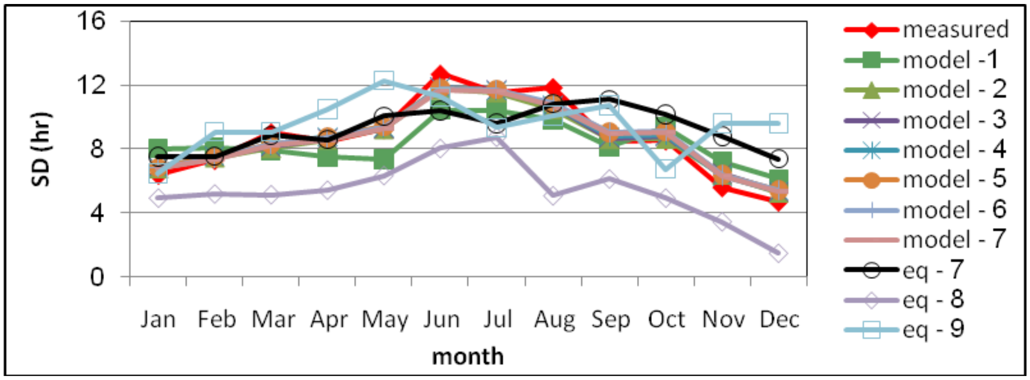


Figure (3). Comparison between the measured and predicted sunshine duration 2001 for Tripoli.

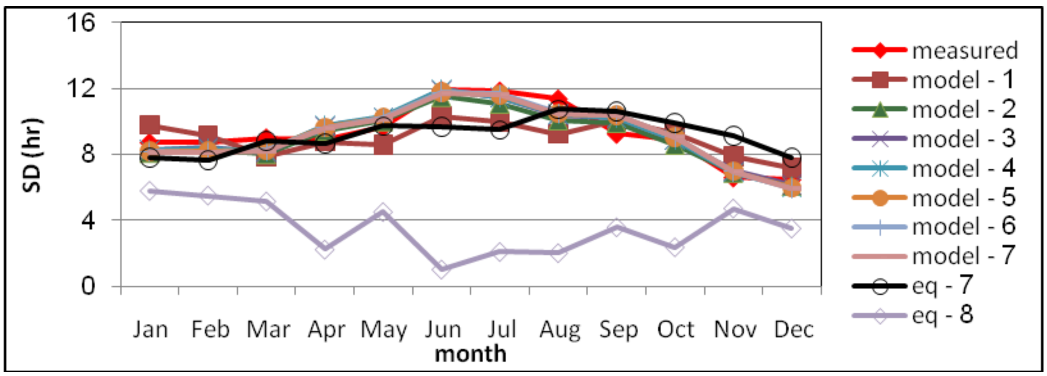


Figure (4). Comparison between the measured and predicted sunshine duration 2001 for Sirt.

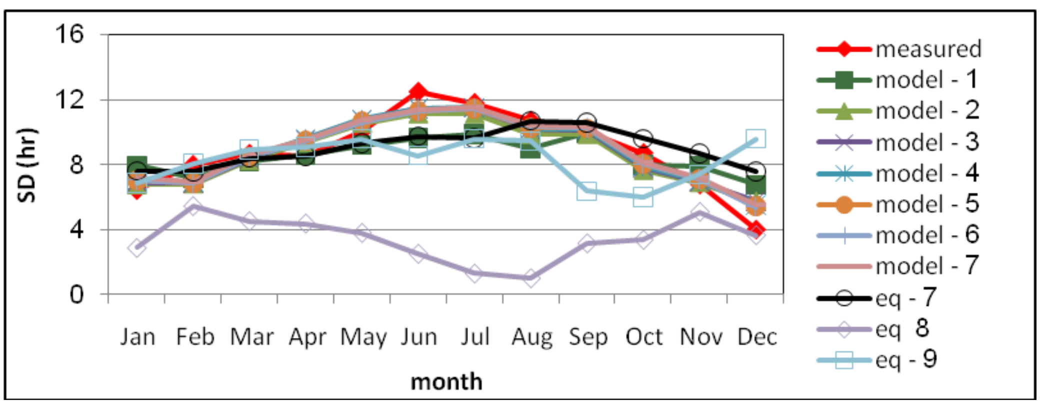


Figure (5). Comparison between the measured and predicted sunshine duration 2001for Tubruk.

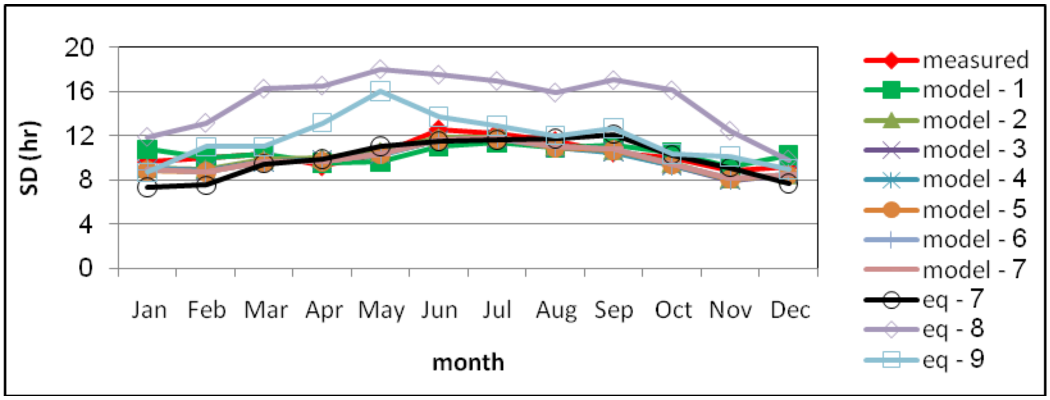


Figure (6). Comparison between the measured and predicted Sunshine duration 2001 for Sabha.

This is owing to the fact that the values of MAE (hr), RMSE (hr), RMSE (%) and NSE (%) of the all models except model (1) were in the ranges of their favorable values. Table (6) shows that the range of MAE (hr), RMSE (hr), RMSE (%) and NSE are from 0.46 h to 0.76 h, 0.59 h to 0.90 h, 6.53 % to 10.12 % and 58.92 % to 93.54%, respectively. However, it is clear that model 2 with two parameters i.e., cloudiness index and total day length shows high capability in predicting sunshine duration as the range of -1.26 to 0.62 with R^2 in the range of 0.231 to 0.975. The lowest values of MAE (hr), RMSE (hr), RMSE (%) are 1.06 h, 1.41 h, 12.93 %, respectively. The values of NSE (%) are weak whereas the best value equals to 42.38. %. However, Equation (7) performed quite well as compared with equations 8 and 9.

The results appear consistent with previous studies in which regression approach were used to determine sunshine duration. Matzarakis and Katsoulis (2006) estimated mean annual and seasonal duration of bright sunshine for Greece from multiple regression equation which depends similar as other models with more than two parameters.

It is apparent from table 4 that models 1 to 7 have better performance in estimating the sunshine duration as compared with Abdelwahed and Snyder (2015) equations. Abdelwahed and Snyder (2015) equations produce poor performance of α, R^2, MAE (hr), RMSE (hr), RMSE (%) and NSE. It was found that the slope values (α) which combined with real values of determination coefficient are in on distance of each station from the nearest cost, height above sea level for each station location, percentage of land cover around each station, latitude of each station, and longitude of each station. It was found that RMSE were found to be9.90 h, 6.15 h, 4.69 h, 6.22 h, and 5.33 h for annual sunshine, winter, spring, summer, and autumn, respectively. El- Metwally (2005) also estimated relative sunshine (n/N) for six sites in Egypt as a function of monthly mean daily maximum and minimum temperatures and cloud cover index. He found that the RMSE% values of the predicted relative SD varied from 2.3% to 14.5%, respectively.

According to the results of the performance indicators analysis, the comparisons between the measured and predicted data by models 2 to 7 indicated high agreement and could be recommended in predicting the sunshine duration under Libyan conditions with high precision. Also, the results revealed that predicting SD depending on only one variable is not safe and not reliable as equations 7, 8 and 9 show.

Table (6) Evaluation parameters for predicting sunshine hours models

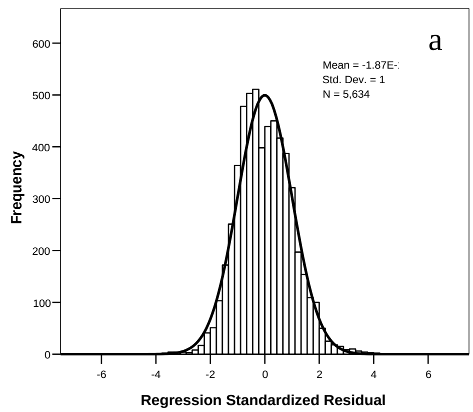
Model	A	R^2	MAE (hr)	RMSE (hr)	RMSE %	NSE
			TRIPOLI AIRPORT			
1	0.930	-0.005	1.33	1.44	16.68	62.63
2	0.974	0.903	0.46	0.63	7.33	92.78
3	0.986	0.907	0.49	0.60	6.93	93.54
4	0.999	0.910	0.50	0.59	6.78	93.83
5	0.995	0.901	0.53	0.61	7.03	93.37
6	0.995	0.900	0.53	0.61	7.07	93.29
7	0.983	0.884	0.49	0.61	7.05	93.32
Eq. 7	1.01	-1.01	1.47	1.79	20.00	42.48
Eq. 8	0.625	0.677	3.24	3.51	40.62	-121.00
Eq.9	1.044	-1.34	2.09	2.48	28.65	-10.28
			SIRT			
1	0.952	0.420	1.05	1.21	12.99	48.73
2	0.968	0.860	0.59	0.66	7.07	84.82
3	0.991	0.867	0.51	0.61	6.53	87.03
4	0.991	0.874	0.52	0.61	6.54	86.99
5	0.993	0.870	0.54	0.62	6.64	86.62
6	0.992	0.869	0.55	0.63	6.73	86.23
7	0.988	0.869	0.56	0.63	6.76	86.13
Eq. 7	0.966	-0.870	1.16	1.41	15.27	29.12
Eq. 8	0.352	-0.640	5.74	6.41	69.00	-134.68
			TUBRUK			
1	0.925	-0.840	1.24	1.52	17.11	55.60
2	0.963	0.787	0.76	0.90	10.12	84.48
3	0.985	0.808	0.71	0.84	9.53	86.24
4	0.990	0.834	0.67	0.79	8.90	88.00
5	0.988	0.838	0.67	0.77	8.70	88.54

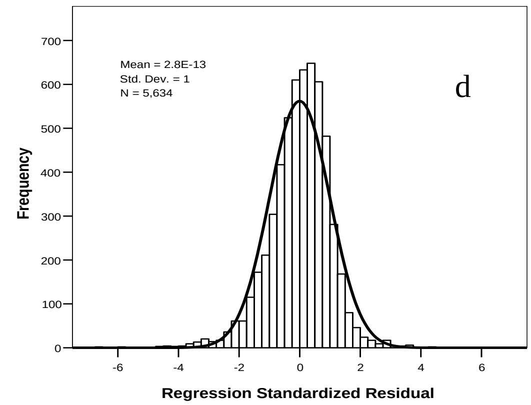
6	0.987	0.842	0.66	0.76	8.64	88.69
7	0.998	0.831	0.66	0.78	8.78	88.33
Eq. 7	-1.260	0.975	1.19	1.61	18.27	49.37
Eq. 8	0.344	-1.10	5.42	6.28	70.96	-663.27
Eq.9	0.880	-2.14	2.34	3.43	38.75	-20.30
SABHA						
1	0.994	-0.230	0.63	0.75	7.26	55.53
2	0.963	0.829	0.58	0.63	6.04	69.30
3	0.951	0.850	0.60	0.68	6.57	63.65
4	0.948	0.848	0.61	0.69	6.68	62.42
5	0.948	0.831	0.64	0.72	6.98	58.92
6	0.952	0.831	0.62	0.69	6.69	62.31
7	0.953	0.829	0.62	0.69	6.66	62.68
Eq. 7	0.959	0.460	1.05	1.67	12.93	-31.06
Eq. 8	1.458	0.379	4.76	5.17	49.92	-200.05
Eq.9	1.120	0.231	1.54	2.21	21.38	-285.39

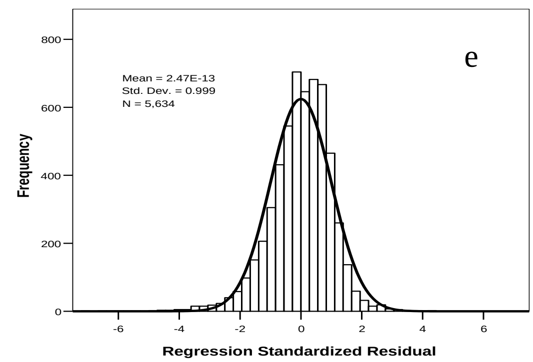
Conclusions

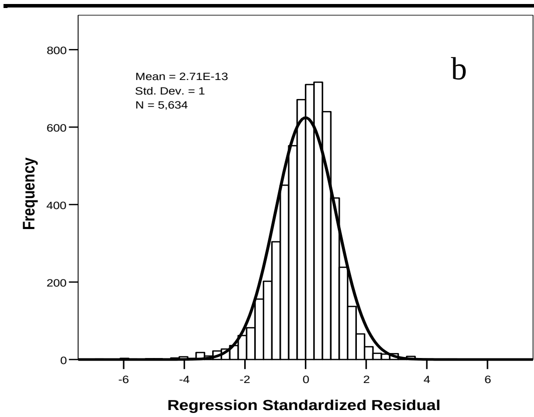
Direct measurements of sunshine duration may be unavailable for some stations and for some period due to spatial, temporal and cost constraints. As an alternative, indirect estimation of sunshine duration using stepwise regression technique was adopted to develop sunshine duration models for Libya. Several comparisons among the developed models were conducted to asses these models. According to determination coefficient (R^2), it was found that the developed models explained 58% to 84% of the variance in the sunshine duration. The statistical evaluation indicators with the highest values of R^2, NSE, least values of MAE, RMSE, RMSE % and slope values (α) close to one showed that six of the developed models were superior to the model with one parameter (cloudiness index) in estimating the sunshine hours. In addition, the accuracy of model 2 with two parameters (cloudiness index and total day length) does not differ too much from the accuracy of the other models with more than two parameters. The results also revealed that the estimations of Abdelwahed and Snyder (2015) equations were worse than that of our models. The advantages of the developed models can be attributed to the large relevant inputs which cover diverse geographically and climatologically zones of Libya. However, the developed equations can be strongly recommended for estimating the missing or unavailable data values of sunshine duration in Libya.

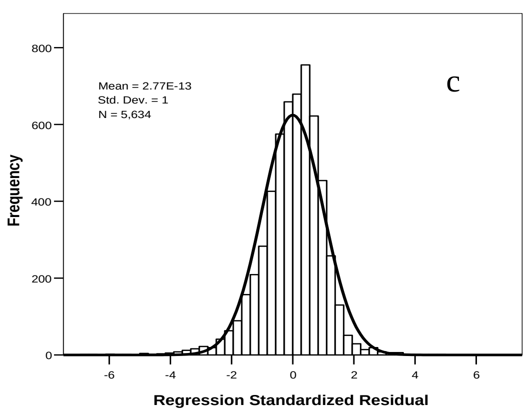
Appindex1: Histogram of Regression Standarized Residual for the Models 1(a), 2(b), 3(c), 4(d), 5(e), 6(f), 7(g).

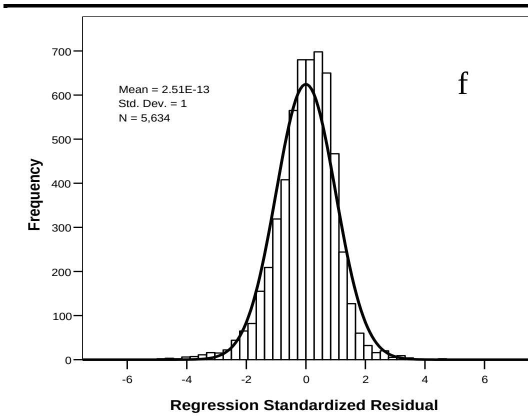

a
Mean = -1.87E-
Std. Dev. = 1
N = 5,634
Frequency
Regression Standardized Residual

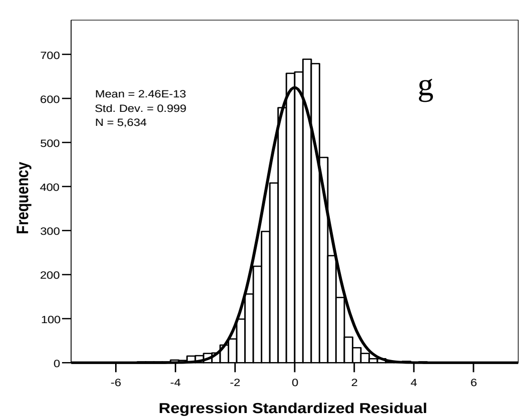

d
Mean = 2.8E-13
Std. Dev. = 1
N = 5,634
Frequency
Regression Standardized Residual


e
Mean = 2.47E-13
Std. Dev. = 0.999
N = 5,634
Frequency
Regression Standardized Residual


b
Mean = 2.71E-13
Std. Dev. = 1
N = 5,634
Frequency
Regression Standardized Residual


c
Mean = 2.77E-13
Std. Dev. = 1
N = 5,634
Frequency
Regression Standardized Residual


f
Mean = 2.51E-13
Std. Dev. = 1
N = 5,634
Frequency
Regression Standardized Residual


g
Mean = 2.46E-13
Std. Dev. = 0.999
N = 5,634
Frequency
Regression Standardized Residual

References:

Abdelwahed M. H. and Snyder R. L. 2015. Calculating sunshine hours and reference evapotranspiration in arid regions when solar radiation data are limited. Irrigation and Drainage journal. 64: 419-425.

Abdul Al-Azzawi. 2006. Photonics: Principles and Practices. CRC Press. Taylor & Francis Group.

Cameron. C., Windmeijer A., Frank ; Gramajo A. G., Cane H. and Khosla D. 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. Journal of Econometrics. 77 (2): 1790 -2.

Duffie, J. A. and Beckman, W. A. 1991. Solar engineering of thermal processes, JohnWiley & Sons, New York, NY, USA.

Ekhmaj, A. I.; Ben zagta M., Ezlit Y. D. and Shaghleb, A. M. 2013. Estimation of the reference evapotranspiration in Sirt region using an artificial neural networks. The Libyan Journal of Agriculture. Vol. (18) No: (1,2): 1- 14.

El-Metwally, M. 2005. Sunshine and global solar radiation estimation at different sites in Egypt. Journal of Atmospheric and Solar-Terrestrial Physics. 67: 1331–1342.

El-Tantawi, A. M. 2005.Climate change in Libya and desertification of Jifara plain: using Geographical Information System and Remote Sensing Techniques, Dissertation submitted in der Johannes Gutenberg- Universitat in Mainz (Germany). 308 pp.

Gueymard, C.A. and Myers, D.R. (2008) Solar radiation measurement: progress in radiometry for improved modeling. In: Badescu V (ed) Modeling solar radiation at the earth surface. Springer, Berlin.

Hinssen, Y. B. L., 2006: Comparison of different methods for the determination of sunshine duration. KNMI Scientific Rep. WR–2006-06, 72 pp.

Kandirmaz, H. M., Kaba, K. and Avci, M. 2014. Estimation of monthly sunshine duration in Turkey using artificial neural networks. International Journal of Photoenergy. Volume 2014, Article ID 680596, 9 pages.

Karume, K.; Banda EJKB; Mubiru,J. and Majaliwa, M. 2007. Correlation between sunshine hours and climatic parameters at four locations in Uganda. Tanzania Journal of Science.Vol: 03. 93- 100.

Khalifeloo, M. H., Mohammad, M. and Heydari, M. 2015. Multiple imputation for hydrological missing data by using a regression method (Klang River Basin). International Journal of Researchin Engineering and Technology. Vol (4): 06 (519-524).

Martina, Z.; Pavol, P. and Zofia, K. 2014. Statistical and geographical analysis of precipitation and temperatures of Libya. Discovery, Volume: 22(72): 37-41.

Martyn, D. 1992: Climates of the World. Elsevier, 436 pp.

Matzarakis, A. P. and Katsoulis, V. P. 2006. Sunshine duration hours over the Greek region.

Theoretical and Applied Climatology. Volume 83, Issue 1, pp 107-120.

MÜHR, B. 2004. Die Klimaklassifikation nach W. Köppen,(German-Edition) http://www.klimadiagramme.de/ Frame/koeppen.html 02.06.2016.

Nash, J. E., and J. V. Sutcliffe.1970. River flow forecasting through conceptual models, 1, A discussion of principles, J. Hydrology. 10. (282-290).

Reddy, S. J. 1974. An empirical method for estimating sunshine from total cloud amount. Solar Energy, 15(4), 281–285.

Robaa, S. M. 2008. Evaluation of sunshine duration from cloud data in Egypt. Energy, vol. 33, no. 5, pp. 785–795.

Sen, Z. and Eljadid, A. G. 1999. Rainfall distribution function of Libya and rainfall prediction. Hydrological Science journal. 44(5): 665- 680.

SPSS, 2003. Sigma Plot for Windows. Ver. 11.5. Chicago, Ill.: SPSS, Inc.

Startz, R. 2013. EView Illustrated for Version 8. HIS Global . Inc. pp. 434.

Tveito, O. E., 1998. Spatial estimation of mean monthly temperatures by multiple linear regressions. DNMI-klima 18/98: 1–19.

Umoh, M. D.; Udo S. O. and Udoakah, Y. N. 2013. Global solar radiation studies due to sunshine hours and other meteorological parameters. American Journal of Engineering and Applied Sciences 6 (3): 320-326.

Velds, C. A. 1992. Solar radiation in the Netherlands. Thieme /KNMI, 166 pp.

WMO, 2008: Guide to Meteorological Instruments and methods of observation. 7th ed., WMO Rep. 8, World Meteorological Organization, Geneva, Switzerland. 681pp.

Yang, Y.H.; Zhao, N. Hao, XF and Li, CQ. 2009. Decreasing trend of sunshine hours and related driving forces in North China. Theor Appl Climatol. 97:91 –98.

Yin, X. 1999. Bright sunshine duration in relation to precipitation, air temperature and geographic location. Theoretical and Applied Climatology. Volume 64, Issue: 1, pp 61-68.

المجلة الليبية للعلوم الزراعية
المجلد (21): العدد (1- 2): 2016: 92-111
ورقة أصيلة

تقدير فترة السطوع الشمسي باستخدام مقاربة إحصائية، ليبيا كحالة دراسية

أحمد إبراهيم خماج و ميلاد عمران الورشفاني

قسم التربة والمياه- كلية الزراعة- جامعة طرابلس

المستخلص

تعتبر فترة السطوع الشمسي مؤشراً أساسياً للغلاف الجوي، والتي تستعمل في العديد من التطبيقات الزراعية و المعمارية والطاقة الشمسية. إن تطوير طرقاً بديلة غير مباشرة تعتمد على أسس نظرية لتحديد فترة السطوع الشمسي ضرورية في العديد من الحالات التي لا تتوفر فيها بيانات لفترة السطوع الشمسي وذلك نتيجة لمحددات زمانية و اقتصادية . تم في هذه الدراسة، تطوير سبع نماذج لتقدير فترة السطوع الشمسي في ليبيا باستخدام تقنية الانحدار المتدرج، حيث تختلف المتغيرات المستخدمة كمدخلات من نموذج إلى آخر وقد تضمنت تلك المدخلات، المتوسطات الشهرية لكل من مؤشر تغطية السحب، طول فترة النهار الكلي، الرطوبة النسبية المتوسطة، عمق الهطول، درجة الحرارة القصوى، بالإضافة إلى الارتفاع عن مستوى سطح البحر وخطي العرض والطول لستة عشر محطة إرصاد مناخية منتشرة في ليبيا، خلال الفترة بين عامي 1961 و 2000. أجريت عملية التقييم للنماذج التي تم تطويرها باستخدام البيانات المتوفرة لأربع محطات أرصاد مناخية تمثل مناطق مختلفة في المناخ و في الطبيعة الجغرافية لسنة 2001. كما تناولت عملية التقييم أيضا المعادلات المطورة من قبل Abdelwahed and Snyder (2015) لتقدير فترة السطوع الشمسي. تضمنت المعايير الإحصائية المستخدمة في عملية التقييم كل من ، متوسط الخطأ المطلق (MAE) و الجذر التربيعي لمتوسط الخطأ(RMSE) و النسبة المئوية للجذر التربيعي لمتوسط الخطأ (RMSE %) و معامل ناش ساتيكليف (NSE)، كما استخدمت في عملية التقييم معادلة الانحدار الخطي التي تربط العلاقة بين البيانات المقدرة والبيانات المقاسة بقيمة خطأ تقديري مساو للصفر، وكذلك معامل التحديد (R^2). ووفقاً لمؤشرات الأداء، فلقد أظهرت النتائج أن النموذج الذي يحتوي على متغير واحد (مؤشر تغطية السحب) أقل دقة في تقدير فترة السطوع الشمسي مقارنة بالنماذج الأخرى. كما وجد أن النماذج المتحصل عليها من هذه الدراسة أعطت أداءً أفضل في تقدير فترة السطوع الشمسي من تلك المعادلات المطورة من قبل (2015) Snyder Abdelwahed and. وعلى كل حال، يعتبر النموذج الثاني الذي يحتوي على متغيرين (مؤشر تغطية السحب و طول فترة النهار الكلي) مرضٍ وكافٍ وبالإمكان استخدامه بكل ثقة لتقدير فترة السطوع الشمسي في ليبيا.

الكلمات الدالة: فترة السطوع الشمسي، الانحدار المتدرج، نموذج إحصائي.

للاتصال: أحمد ابراهيم خماج. قسم التربة والمياه- كلية الزراعة- جامعة طرابلس - طرابلس - ليبيا.
هاتف: 926577440 218+ **البريد الالكتروني:** *khmaj1@yahoo.com*

استلمت بتاريخ: 2016/9/20 **أجيزت بتاريخ: 2016/12/1**