

Estimation of Global Radiation using Angstrom-Type Models at Selected Sites in Kenya

Ronno, C.K.

Department of Physics, University of Eldoret, P.O. Box 1125 Eldoret-30100, Kenya
Email: ronno2132@gmail.com

Abstract

The 1970s energy crisis stirred a world-wide interest in, and development of, renewable and sustainable energy including solar energy. Solar radiation data is an important input component in the design and implementation of solar energy systems. Only a few scattered meteorological stations in Kenya measure solar radiation on a continuous basis. The use of models to estimate solar radiation is one way to get around this problem. The Angstrom-Type Model is obtained from measured global solar radiation on a horizontal surface situated on Earth's surface. Using solar radiation and duration of sunshine data from 11 Kenyan meteorological stations, this study tested seven Angstrom-Prescott type regression models for their suitability to estimate clearness index for normal skies (K_0). The Angstrom-Prescott type models were obtained by regressing sunshine duration against clearness index and obtaining curves of best fit. Linear, quadratic, exponential, power and logarithmic fits were obtained. Model performance was measured using goodness of fit statistics. These included Pearson correlation coefficient (r), coefficient of determination (R^2), Mean Bias Error (MBE), Root Mean Square Error (RMSE), Student t-statistic, and the t-test. Out of the 11 stations considered only data from Dagoretti, Eldoret, JKI Airport, and Voi meteorological stations showed high R^2 values and these were used to produce modified Angstrom-Prescott models whose long-term, short-term, and overall performances were measured using MBE, RMSE, and t-statistic respectively. For each of the four stations with high R^2 values 10 pairs of equations, one each for (K_0) are presented. These equations maybe used in estimating clearness indices at the four stations using measured fraction of duration of sunshine as the only input. The correlation coefficients for the linear model of the Angstrom type for JKI Airport and Eldoret were 0.274, 0.354 and 0.309, 0.465 respectively while those for the quadratic model for the same two stations were respectively 0.250, 0.430, -0.075 and 0.210, 0.800, -0.270. Recommendations on the use of these models is all models derived from data from Dagoretti, Eldoret, JK Airport, and Voi and expressed in equation form (Eqn. 7 to Eqn. 22) may be used to estimate the clearness indices K_0 for sites at or near the four locations. Equations 23 to 26, pertaining to the Logarithmic model must be used with caution as they do not guarantee best short-term or long-term performance in any of the stations.

Keywords: Global Radiation, Angstrom-Prescott, Regression, Sunshine duration.

INTRODUCTION

The design and implementation of solar energy systems requires detailed knowledge of solar radiation received on a surface situated at the project site. Solar energy systems are to be found in diverse applications such as architecture, agriculture, lighting, cooling and air-conditioning.. In all these applications solar radiation is a key input. A solar energy system designer needs to know the temporal and spatial variation of solar radiation for purposes of sizing and modelling of the system for optimum performance. For non-tracking solar photovoltaic and heating applications it is imperative to have high quality global and diffuse solar radiation falling on the plane-of-array of the collector. A tracking system has only direct solar radiation as input. Concentrating systems are of necessity tracking and would thus require knowledge of direct solar radiation. But of greatest practical value is continuous long-term solar radiation data measured on daily, hourly and sub-hourly basis. Such data is rarely available in a developing country like Kenya due to prohibitive cost of instruments. Historically such impediments have led to the development of ways of estimating, modelling and predicting solar radiation at any given location. Most such approaches are based on correlations between solar radiation and meteorological and geographical parameters such as sunshine duration, cloud cover, atmospheric turbidity, ambient temperature, altitude, latitude and relative humidity on the one hand and modelling of optical-atmospheric interaction on the other.

The demand for solar resource data is on the increase due to the need to make renewables part of the world's energy mix. Renewables are greener and environmentally friendlier than non-renewable sources of energy. Apart from the Kyoto Protocol's requirement that nations reduce their Greenhouse Gas Emissions (GHE), in Kenya there is a statutory imperative in the form of the following legislations: The Energy Regulatory Commission (ERC, 2012) requires large energy consumers and producers to comply with the Energy Management Regulations, 2012 and in addition it behooves them to invest in energy efficiency and conservation (GOK, 2014). The National Environmental Management Authority (NEMA) imposes sustainability parameters on energy producers and consumers resulting in a shift towards renewable energy technologies. Lastly, the National Building Authority Regulations 2013 of the National Construction Authority (NCA) prescribes that certain categories of residential and commercial buildings must incorporate solar heating as a portion of its hot water provisions. These statutes will guarantee an increased demand for raw and processed solar radiation data as it forms the starting point for solar energy system design.

The aim of this study was to develop models that estimate the long-term mean monthly global irradiance over Kenya using Angstrom-Type model. To achieve this the following objectives were set out. One: establish the nature of correlation between global solar radiation and sunshine duration. Two: Fit the data into various models based on regression of fractional sunshine duration ($\bar{S}/\bar{S}_0$) and normal-sky clearness index ($K_0 = \bar{H}/\bar{H}_0$) so as to identify the best models to estimate global radiation. Three: Evaluate the reliability and usability of the models.

MATERIALS AND METHODS

The irradiance and sunshine data used in this study was retrieved from World Radiation Data Centre's (WRDC) data servers at the National Renewable Energy Laboratory (NREL) Golden, Colorado, USA, via online protocols. Detailed instructions for submitting data files are found in Krichak (1987), a WMO document that also serves as a comprehensive booklet for those using the data.

Data available from Kenya's Meteorological Service consists of two parameters, namely daily sums and monthly mean of daily sums of global solar radiation in units of 0.01 MJ/m^2 or J/cm^2 and, the monthly sum, as well as monthly mean of daily sums of bright sunshine duration in units of 0.1 hr. For the purpose of this study, the meteorological stations selected were those with the longest possible record and also with valid sets of data, namely containing both solar radiation and sunshine duration for the same month. These stations were, Dagoretti (DAG), Eldoret (ELD), Garissa (GAR), JKI Airport (JKIA), Kisumu (KIS), Kitale (KIT), Lamu (LAM), Lodwar (LOD), Malindi (MAL), Moi International Airport Mombasa (MIA), and Voi (VOI). The WRDC radiation data record for Kenya covers various periods with the earliest going back to 1964. However, the majority of sunshine data go back to 1969. The WRDC record has been compiled up to June 1993 and there has been no update since.

For each station, continuous records of complete pairs of the two quantities were extracted for the entire period covered. The monthly mean of daily clearness index $K_0 = \bar{H} / \bar{H}_0$ and the monthly mean of daily fraction of bright sunshine $\bar{S} / \bar{S}_0$ were then calculated and used in various regressions and calculations.

For each station, the clearness index (K_0) data was regressed against fraction of bright sunshine duration and curves of best fit obtained. The result of the regression is Angstrom-type models. An overview of the development of the Angstrom-Prescott-Page equation is found in Martinez-Lozano, Tena, Onrubia, and Rubia (1984). The linear Angstrom-type models have the form:

$$\bar{H} / \bar{H}_0 = a + b(\bar{S} / \bar{S}_0) \quad (1)$$

Here $\bar{H}$ is the measured monthly mean of daily global (total) radiation falling on a horizontal collector located on Earth's surface while $\bar{H}_0$ is the corresponding value for a surface located at the top of the atmosphere at same latitude and longitude. Furthermore $\bar{S}$ is the measured monthly mean of daily bright sunshine duration while $\bar{S}_0$ is the corresponding monthly mean of daily daylength (time from sunrise to sunset). The regression constants a and b are in general site dependent (Baker and Haines, 1969; Panoras and Mavroudis, 1994).

Recent work of this type include for example Gadiwala, Usman, Akhtar, and Jamil (2013) as well as Razmjoo, Heibati, Ghadimi, Qolipour and Nasab (2016). Quadratic, exponential, power, and logarithmic extensions of the Angstrom-type model were also used in this work.

The resulting five models were subjected to the following goodness of fit statistics:

Pearson Correlation Coefficient: measures the strength of linear dependence between two variables, and is given by,

$$r_{xy} = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{(N \sum x_i^2 - (\sum x_i)^2)(N \sum y_i^2 - (\sum y_i)^2)}} \quad (2)$$

Here x is the independent variable (fraction of bright sunshine $\bar{S}/\bar{S}_0$) and y is the dependent variable (Clearness index $\bar{H}/\bar{H}_0$). This parameter detects the nature and magnitude of the relationship between the measured values and those predicted by the model.

Coefficient of Determination R^2 : This is the square of the Pearson coefficient (commonly called R-Squared). It measures the success of the regression in predicting the values of the dependent variable within the sample. $R^2 = 1$ if the regression fits perfectly, and zero if it fits no better than the sample mean of the dependent variable. R^2 is the fraction of the variance of the dependent variable explained by the independent variables. It can be negative if the regression does not have an intercept, or constant, if two-stage least squares is used.

The Root Mean Square Error (RMSE): Is defined as

$$RMSE = \left(\frac{1}{N} \sum_{i=1}^N d_i^2 \right)^{\frac{1}{2}} \quad (3)$$

Here N is the number of data pairs and d_i is the difference between i^{th} estimated and i^{th} measured values. This test provides information on the short-term performance of the model by comparing measured and estimated values term by term. A smaller RMSE indicates better model performance. However, a few large errors in the sum can produce an undesirable increase in the RMSE.

The Mean Bias Error (MBE): Is given by

$$MBE = \left(\frac{1}{N} \sum_{i=1}^N d_i \right) \quad (4)$$

This test measures the long-term performance of a model. A low MBE is desired. A positive value gives the average amount of overestimation in the calculated value and vice-versa. A shortcoming of this test is that overestimation of an individual observation will cancel underestimation in another observation.

The t-statistic: Allows the comparison between the performances of various models and in addition may be used to calculate the model's statistical significance to any stated confidence level. In terms of the MBE and RMSE the t -statistic for N data points, is given as:

$$t_{stat} = \left(\frac{(N-1)MBE^2}{RMSE^2 - MBE^2} \right)^{0.5} \quad (5)$$

The smaller the value of t -statistic, the better the overall performance of the model.

The Student's t-Distribution: The statistical significance of the correlation coefficient was determined using the 2-tailed test of Student's t distribution given by

$$t = r_{xy} \sqrt{\frac{N-2}{1-R^2}} \quad (6)$$

RESULTS AND DISCUSSION

Stations with high correlation ($r > .800$) between clearness index and fraction of bright sunshine were chosen for further consideration. These were JKI Airport, Eldoret, Dagoretti, and Voi.

Table 1: Goodness of Fit Statistics for Stations with High Coefficient of Determination

MODEL	STATION	MBE	RMSE	t-statistic
LINEAR	Dagoretti	-.00013	.048	.038
	Eldoret	-.00018	.026	.109
	JK Airport	.00020	.029	.066
	Voi	-.00007	.029	.028
QUADRATIC	Dagoretti	.00044	.048	.129
	Eldoret	.00017	.026	.100
	JK Airport	.00031	.029	.102
	Voi	.00014	.029	.055
EXPONENTIAL	Dagoretti	-.00015	.048	.044
	Eldoret	-.00037	.027	.212
	JK Airport	-.00008	.029	.025
	Voi	-.00009	.029	.035
POWER	Dagoretti	.00001	.048	.002
	Eldoret	.00015	.026	.091
	JK Airport	.00000	.029	.001
	Voi	.00024	.029	.091
LOGARITHMIC	Dagoretti	.00027	.048	.080
	Eldoret	.00000	.026	.000
	JK Airport	.00015	.029	.051
	Voi	.00028	.029	.115

Linear Model: For a true linear relationship, a high R^2 is desirable. This condition holds for Dagoretti ($R^2 = .728$), Eldoret ($R^2 = .784$), J K Airport ($R^2 = .819$), and Voi ($R^2 = .755$). This means there is little difference between measured and estimated values for these stations. The resulting model equations for the stations with high R^2 values are given in a later section below. These equations may be used to estimate clearness index K_0 given the corresponding input sunshine fraction $\bar{S} / \bar{S}_0$. A low MBE is desirable and is an indication of long-term performance of the model. Three stations returned very low but negative MBEs indicating underestimation in their long-term model performance while JKA had a low positive (MBE = .00020) MBE making its linear model the best long-term one. A low RMSE value is desirable for good short-term model performance. Eldoret had the lowest (RMSE = .026) followed by both J K Airport and Voi, with Dagoretti (MBE = .048) having the highest. This makes Eldoret the station with the best short-term performance for Linear Model. A low t-statistic indicates overall model performance making the linear model for Voi ($t\text{-stat} = .028$) the overall best one.

Quadratic Model: Voi had the best long-term (lowest MBE) quadratic model (MBE = .00014) while the best short-term model was obtained for Eldoret (lowest RMSE =

.026). The overall best performing quadratic model (lowest t-statistic) corresponds to Voi ($t\text{-stat} = .055$).

Exponential Model: The MBE test shows that models for all four stations underestimated the calculated value (negative MBE). JKI Airport had the absolute minimum MBE (MBE = .00008) thus making it the station with the best long-term performing model. The lowest RMSE was obtained for Eldoret (RMSE = .027) making it the best performing short-term performing Exponential Model. The t-statistic test indicated that JKI Airport had the minimum value ($t\text{-stat} = .025$) making it the best overall performing Exponential Model.

Power Model: The Power Model for JKI Airport had the lowest MBE (MBE = .00000) making it the station with the best long-term performing Power Model. Eldoret returned the lowest RMSE value (RMSE = .026) making it the best short-term performing Power Model. The best overall performing Power Model was noted to be JKI Airport ($t\text{-stat} = .001$).

Logarithmic Model: Results of the MBE and RMSE tests showed that Eldoret, has lowest MBE (MBE = .00000) RMSE (RMSE = .026) meaning is the best performing long-term and short-term Logarithmic Model. This is a contradiction. The next lowest MBE (MBE = .00015) is obtained for JKI Airport which implies the lowest RMSE must be that of Voi (RMSE = .029). But this is also the RMSE value for JKI Airport. The logical conclusion is that the MBE and RMSE tests cannot discriminate between the long and short term performance of the Logarithmic Model. The t-statistic however indicates that Eldoret ($t\text{-stat} = .000$) has the best overall performing Logarithmic Model.

Model Equations

The equations for the Linear, Quadratic, Exponential, Power, and Logarithmic Angstrom-Type Models are given below

Linear Models

$$\text{Dagoretti: } K_0 = 0.234 + 0.505(\bar{S} / \bar{S}_0). \quad (7)$$

$$\text{Eldoret: } K_0 = 0.039 + 0.465(\bar{S} / \bar{S}_0). \quad (8)$$

$$\text{JKI Airport: } K_0 = 0.274 + 0.354(\bar{S} / \bar{S}_0). \quad (9)$$

$$\text{Voi: } K_0 = 0.219 + 0.533(\bar{S} / \bar{S}_0). \quad (10)$$

Quadratic Models

$$\text{Dagoretti } K_0 = 0.68 - 0.90(\bar{S} / \bar{S}_0) + 0.48(\bar{S} / \bar{S}_0)^2. \quad (11)$$

$$\text{Eldoret \& } K_0 = .21 + 0.80(\bar{S} / \bar{S}_0) - 0.27(\bar{S} / \bar{S}_0)^2. \quad (12)$$

$$\text{J K Airport } K_0 = .025 + 0.43(\bar{S} / \bar{S}_0) - 0.075(\bar{S} / \bar{S}_0)^2. \quad (13)$$

$$\text{Voi } K_0 = 0.11 + 0.90(\bar{S} / \bar{S}_0) - 0.30(\bar{S} / \bar{S}_0)^2. \quad (14)$$

Exponential Models

$$\text{Dagoretti } K_0 = 0.293 \exp[0.994(\bar{S} / \bar{S}_0)]. \quad (15)$$

$$\text{Eldoret } K_0 = 0.369 \exp[0.771(\bar{S} / \bar{S}_0)]. \quad (16)$$

$$\text{J K Airport} \quad K_0 = 0.306 \exp[0.757(\bar{S} / \bar{S}_0)] \quad (17)$$

$$\text{Voi:} \quad K_0 = 0.297 \exp[0.983(\bar{S} / \bar{S}_0)] \quad (18)$$

Power Models

$$\text{Dagoretti} \quad K_0 = 0.694(\bar{S} / \bar{S}_0)^{0.493} \quad (19)$$

$$\text{Eldoret} \quad K_0 = 0.752(\bar{S} / \bar{S}_0)^{0.473} \quad (20)$$

$$\text{J K Airport} \quad K_0 = 0.592(\bar{S} / \bar{S}_0)^{0.372} \quad (21)$$

$$\text{Voi} \quad K_c = 0.970(\bar{S} / \bar{S}_0)^{0.565} \quad (22)$$

Logarithmic Models

$$\text{Dagoretti} \quad K_0 = 0.670 + 0.564 \log(\bar{S} / \bar{S}_0) \quad (23)$$

$$\text{Eldoret} \quad K_0 = 0.734 + 0.637 \log(\bar{S} / \bar{S}_0) \quad (24)$$

$$\text{JKI Airport} \quad K_0 = 0.578 + 0.387 \log(\bar{S} / \bar{S}_0) \quad (25)$$

$$\text{Voi} \quad K_0 = 0.702 + 0.715 \log(\bar{S} / \bar{S}_0) \quad (26)$$

CONCLUSIONS

The overall aim of this study was to estimate long-term mean monthly global irradiance over Kenya using Angstrom-Type irradiance models. Data from 11 meteorological stations in Kenya were used to generate a set of clearness indices which were then correlated with corresponding sunshine fraction. The Pearson correlation coefficient test indicated that there was a positive and significant correlation between clearness index and fraction of sunshine duration. The strength of the correlation was measured using the coefficient of correlation R^2 and data from J K Airport, Eldoret, Voi, and Dagoretti stations showed high correlation.

The monthly mean of daily global irradiance was then regressed against the monthly mean of daily fraction of sunshine duration for the four stations for each of the 5 models, namely Linear, Quadratic, Exponential, Power and Logarithmic.

Based on the MBE, RMSE and t-statistic respectively stations with best long-term, best short-term and best overall performance were identified for the 5 models. For the Linear model these were JKI Airport, Eldoret and Voi. For the Quadratic model they were Voi, Eldoret and Voi. As for the Exponential model they were JKI Airport, Eldoret and JKI Airport. The Power model JKI Airport, Eldoret and JKI Airport respectively. The MBE and RMSE tests for the Logarithmic model were inconclusive even though the t-statistic seemed to suggest Eldoret had the best overall performing model.

RECOMMENDATIONS

Based on the findings of this study and the conclusions outlined above the following recommendations can be made:

- i. All models derived from data from Dagoretti, Eldoret, JKI Airport, and Voi and expressed in equation form (Eqn. 7 to Eqn. 22) may be used to estimate the clearness indices K_0 for sites at or near the four locations. Equations 23 to 26, pertaining to the Logarithmic model must be used with caution as they do not guarantee best short-term or long-term performance in any of the stations.
- ii. It is also important that a vigorous campaign of solar data measurement and monitoring be initiated throughout Kenya if the solar energy exploitation is to be made full use of. The challenge is for individuals, research institutions and relevant government entities to come together and forge a way forward.

REFERENCES

- Baker, D. G., and Haines, D. A. (1969). Solar Radiation and Sunshine Duration Relationship in the North-Central Region of Alaska. Minnesota Agricultural Experimental Station. Tech. Bull. 262
- ERC. (2012). The Energy (Solar Water Heating) Regulations 2012. Energy Regulatory Commission of Kenya. Government Printing Press.
- Gadiwala, M.S., Usman, A., Akhtar, M., and Jamil, K. (2013). Empirical Models for the Estimation of Global Solar Radiation with Sunshine Hours on Horizontal Surface in Various Cities of Pakistan. *Pakistan Journal of Meteorology*. 9, 43.
- GOK. (2014). Draft National Energy Policy. Government of Kenya Ministry of Energy and Petroleum. Government Printing Office
- Hottel, H. C. (1976). A simple model for estimating the transmittance of direct solar radiation through clear atmospheres. *Solar Energy*. 18, 129.
- Krichak, M. O. (1987). Input Format Guidelines for World Radiometric Network Data. Leningrad. Downloadable from <http://wrdc-mgo.nrel.gov/html/formats.html>
- Martinez-Lozano, J. A., Tena, F., Onrubia, J. E., and De la Rubia, J. (1984). The Historical Evolution of the Angstrom formula and its Modifications: Review and Bibliography. *Agricultural and Forest Meteorology*. 33.
- Panoras, A. G., and Mavroudis, I. G. (1994). Relation between duration of sunshine and global solar radiation in Thessaloniki Plain. *Hydrotechnica: J. Hellenic Hydrotechnical Association*. 4 (1), 3-15.
- Razmjoo, A., Heibati, S. M., Ghadimi, M., Qolipour, M., and Nasab, J.R. (2016). Using Angstrom-Prescott (A-P) Method for Estimating Monthly Global Solar Radiation in Kashan. *Journal of Fundamentals of Renewable Energy and Applications*. 6, 214.