MODELLING OF SOLAR RADIATION IN BALEYSA STATE, NIGERIA USING THE KERNEL REGRESSION METHOD

ARONU, C. O.[1], UKA, C. O.[2], EKWUEME, O. G.[3]AND SOL-AKUBUDE V. I. P[4]
[1]Department of Statistics, Chukwuemeka Odumegwu Ojukwu University, Anambra State
[2]Department of Statistics, Federal College of Agriculture, Ishiagu, Ebony State, Nigeria
[3]Department of Industrial Production Engineering, Nnamdi Azikiwe University, Awka, Nigeria
[4]Department of Statistics, Chukwuemeka Odumegwu Ojukwu University, Anambra State

ABSTRACT
This study focused on modelling total solar radiation in Bayelsa State using the Kernel regression analysis. The objective of the study is to obtain an appropriate Kernel regression model for estimating total solar radiation in Bayelsa State and to determine whether time has a significant impact of total solar radiation. The source of data used for this study was the secondary data obtained from National Aeronautics and Space Administration (NASA). The data comprises of solar radiation values measured over a period of eleven years (2007-2017). The findings of the study found R-square value of 0.983 (98.3%), and bandwidth of 0.1970. The result of the R-square value indicates a strong adequacy of the model which implies that the model is useful for estimating total solar radiation. Also, this result implies that time (Month) was able to explain about 98.3% of total variation in total solar radiation. In addition, findings showed that time does not significantly impact on total solar radiation since the model obtained a p-value of 1.00 which falls on the acceptance region of the hypothesis assuming 95% confidence level.

Keywords: Bandwidth, Solar Radiation, Kernel Regression, Modelling,

1. INTRODUCTION

Solar radiation are useful in providing energy required for photosynthesis and transpiration of crops. Solar radiation has been identified as one of the meteorological factors determining potential yields. Adequate data on the availability of solar radiation on horizontal surface is essential for the optimum design and study of solar energy conversion system. Due to the usefulness and potentials of solar radiation in providing energy for Earth's climatic system, researchers have focused on examining variation that exist in global solar radiation. In areas where radiation measurements are scanty, theoretical forecast of the availability of solar energy can be used to predict these measurements from standard weather parameters that are extensively measured (Abdullahi and Singh, 2014).

Abiodun and Anisiji (2019) noted that planning and development of local solar energy system relies on the availability of adequate solar radiation data which can be used to assess the probable long term solar systems performance and economic viability. The solar radiation received at the surface of the earth varies daily, seasonally and annually depending on the climatic condition of the region. Hence, the need for many years of measurement of solar radiation observations in order to ensure accurate estimation of the parameters and distribution of solar radiation. Also, there is need to develop an appropriate and useful model for the estimation of solar radiation in Nigeria. Several models for estimating solar radiation exists with majority of the authors claiming that the logistic model is the best model for estimating or predicting solar radiation (Temitope, 2015; Abiodun and Anisiji, 2019). It was observed that majority of the researchers in the area of solar radiation have often employ the parametric methods such as analysis of variance (ANOVA), linear regression analysis and factorial design analysis in evaluating solar radiation data without evidence of the data obeying the strict assumptions of the classical methods (Abdulazeez et al., 2010; Abdullahi and Singh, 2014; Adejumo et al., 2017).Hence, the need to use the Kernel regression method to estimate solar radiation in Bayelsa State, Nigeria.

2. LITERATURE REVIEW

Abdulazeez et al. (2010) employed the factorial design in modelling solar radiation data measured from eight different locations in Nigeria. Abiodun and Anisiji (2019) argued that solar radiations have been estimated for some part of Nigeria using different empirical and distribution methods but there exist little literature on solar radiation of South-Southern Nigeria using probability distribution functions.

Abdullahi and Singh (2014) investigated the applicability of solar energy utilization using monthly average daily values of global solar radiation and sunshine hour for five stations of the North-Eastern Nigeria. The findings showed that the clearness index K_T values indicates the clear sky in the months of February to July and October to December for Yola (lat.9.14^{0}N), February to July and September to November for Potiskum (lat.11.42^{0}N), February to July and September to December for Maiduguri (lat.11.51^{0}N), February to July and October to December for Ibi (lat.8.11^{0}N) and February to July and October to December for Bauchi (lat.10.17^{0}N) of the North-Eastern Nigeria. The findings revealed that the solar energy utilization has bright prospects in the North-Eastern Nigeria. The estimated values of global radiation were found to be efficient in compensating for energy deficits. Also, it was found that with the exception of monsoon months August and September, solar energy can be utilized throughout the year for North-Eastern region of Nigeria. In addition, it was found that the maximum values of global solar radiation appear in March, April and May respectively during dry season while minimum values were observed in August and September, respectively during wet season.

Temitope (2015) focused on obtaining the best probable probability distribution model for modelling solar radiation in Ibadan, South western Nigeria. The findings of his study showed that logistic distribution function performs better than the other models considered.

Adejumo et al. (2017) examined the solar radiation readings from three meteorological sites in Nigeria using the Analysis of Variance test (ANOVA). The sites considered were Port Harcourt, Sokoto and Ibadan. The study used data obtained from Nigerian Meteorological Agency (NIMET), Oshodi, Lagos, Nigeria. The findings of the study showed that there exists significant difference in solar radiation in the three sites.

Abiodun and Anisiji (2019) examined the performance of the most suitable statistical distribution function for modelling solar radiation over Yenagoa, Bayelsa State in Nigeria. They employed data obtained from National Aeronautics and Space Administration (NASA). Their study considered six probability distribution functions which comprises of the Normal distribution, Gamma distribution, Lognormal distribution, Logistic distribution, Extreme value distribution, and Weibull distribution. The findings of their study revealed that the logistic distribution is adequate for estimating solar radiation in Bayelsa State.

3. METHODOLOGY

3.1 Source of Data Collection

The source of data used for this study was the secondary data obtained from National Aeronautics and Space Administration (NASA). The data comprises of solar radiation values measured over a period of eleven years (2007-2017).

3.2 Kernel Regression Analysis

The Kernel regression is a non-parametric regression approach which measures the conditional expectation of a random variable (the response variable) (Uka et al., 2018). The major function of the Kernel regression is to obtain a non-linear relation between a pair of random variables the independent variable(s) say X and the dependent variable say Y.

The non-parametric regression model measures the functional relation between Y and X. This implies that the conditional expectation of a variable Y relative to a variable X can be expressed as

$$\mathrm{E}(Y \mid X) = m(X) \tag{1}$$

Where, m is an unknown function to be estimated

The relationship that exist in the Kernel model can be modeled as:

$$Y_i = m(x_i) + \varepsilon_i \tag{2}$$

for i= 1,...,N and $\mathrm{E}(\varepsilon_i) = 0$

where, Y_i and ε_i is n x 1 random vectors called the response vector and the random error vector respectively.

3.3 Kernel Smoothing Techniques

The unknown function m can be estimated using a locally weighted average by using the kernel as a weighing function (Nadaraya, 1964). The weight sequence in the local averaging method can represent the weights distribution by a density function which contains a scale parameter. The scale parameter is used to adjust the size and the form of the weights according to the location of the point with respect to the point of estimation x which is the independent variable (Nadaraya, 1964). This density function is known as the kernel function k. smoothing techniques based on this kind of weight representation are called kernel smoothing. The kernel estimate m(x), is defined as a weighted average of the response variable in a fixed neighborhood around x, determined in shape by the kernel function k and bandwidth h. The kernel used in this study is the Nadaraya-Watson estimator. It is interesting to note that the shape of the kernel weights is determined by the kernel function k with the smoothing parameter h, which is called the bandwidth (Uka et al., 2018; Adjekukor, 2019). The kernel function is a continuous, bounded and symmetric real function which integrates to one.

3.4 Model Specification

The model for this study is expressed as: Equation (3) was obtained by replacing Y_i= solar radiation and x_i = time (months) in equation (2)

$$\mathrm{TSR} = m(\mathrm{Month}_i) + \varepsilon_i \tag{3}$$

for i= 1,...,n where, TSR represents the response variable known as total solar radiation Month represents the independent variable ε_i is n x 1 random vectors called the random error vector

The decision rule is to reject the null hypothesis that time does not impact on total solar radiation when the p-value is less than α=0.05, accept otherwise.

3.5 Data Presentation

Table 1: Table showing the total, mean and Standard Deviation of solar radiation form 2007-2017 in Bayelsa State, Nigeria

Month	Total Solar Radiation (KW-hr/m^2/month)	Mean Solar Radiation (KW-hr/m^2/month)	Standard Deviation of Solar Radiation (KW-hr/m^2/month)
JAN	10.99	0.999091	1.166023
FEB	20.49	1.862727	1.241274
MAR	38.55	3.504545	2.170679
APR	60.07	5.460909	2.030362
MAY	69.02	6.274545	1.328814
JUN	74.86	6.805455	0.89897
JUL	65.15	5.922727	1.483099
AUG	56.53	5.139091	1.79685
SEP	55.16	5.014545	1.828072
OCT	72.86	6.623636	0.989113
NOV	44.39	4.035455	1.871627
DEC	16.86	1.532727	1.835255

Source: National Aeronautics and Space Administration (NASA), 2017

4. DATA ANALYSIS AND RESULT

This section deals with data analysis and result of the analysis.

4.1 Result of Kernel Regression Analysis for estimating Solar radiation

Regression Data: 12 training points, in 1 variable(s)
Month
Bandwidth(s): 0.197035
Kernel Regression Estimator: Local-Constant
Bandwidth Type: Fixed
Residual standard error: 4.594594e-05
R-squared: 0.98332
Continuous Kernel Type: Second-Order Gaussian
No. Continuous Explanatory Vars.: 1

The result of the kernel regression analysis found a residual standard error value of 4.59 x 10^-05 KW-hr/m^2/month, R-square value of 0.983 (98.3%), and bandwidth of 0.1970. The result of the R-square value indicates a strong adequacy of the model which implies that the model is useful for estimating total solar radiation. Also, this result implies that time (Month) was able to explain about 98.3% of total variation in total solar radiation.

4.2 Kernel Regression Significance Test

Type I Test with IID Bootstrap (399 replications, Pivot = TRUE, joint = FALSE)
Explanatory variables tested for significance:
Month (1)
Month
Bandwidth(s): 0.197035
Individual Significance Tests
P Value:
Month 1.00
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The result found a p-value of 1.00 which falls on the acceptance region of the hypothesis assuming 95% confidence level. This result indicates that time (Month) does not significantly impact on total solar radiation in Bayelsa State, Nigeria.

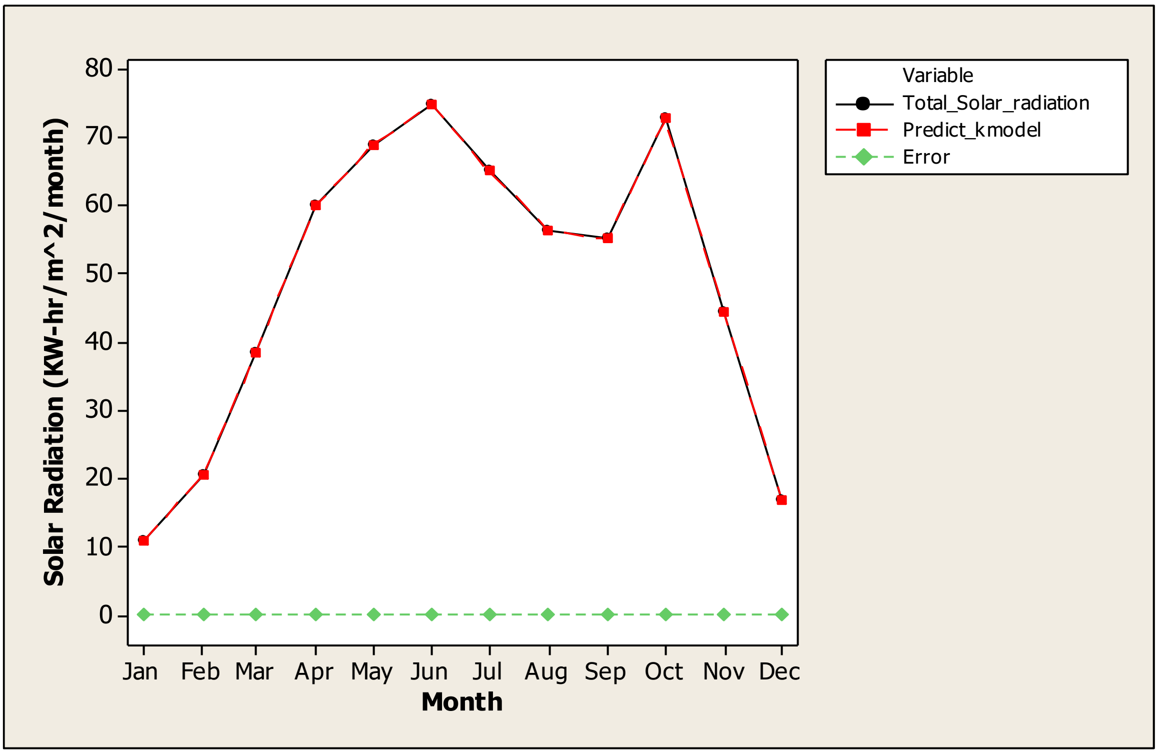

Figure 1: Graph showing the observed total solar radiation, Predicted solar radiation and error

The result displayed in figure 1 validates the claim that the obtained kernel regression model is adequate in estimating total solar radiation since the error is very minimal and the observed (total solar radiation) and the predicted has the same trend.

5. CONCLUSION

This study focused on modelling solar radiation using the Kernel regression analysis. The findings from literature review showed that majority of the researcher in the area of solar radiation have often employ the parametric methods such as analysis of variance (ANOVA), linear regression analysis and factorial design analysis in evaluating solar radiation data without evidence of the data obeying the strict assumptions of the classical methods.

Another approach which does not require data to obey these strict assumptions is the use of the non-parametric methods. The Kernel regression analysis is a non-parametric tool which does not require the data for the analysis to obey the strict classical assumptions such as normality test, homogeneity and test for constant variance.

The findings of the study revealed that the kernel regression model obtained R-square value of 0.983 (98.3%) which was found to be adequate for estimating total solar radiation. Also, the bandwidths for the kernel regression model was found to be 0.1970 which implies that independent variable (Month) has a positive relationship with total solar radiation.

REFERENCES

Abdulazeez, M. A., Ahmed, A., Burari, F. W. (2010). The use of factorial design in the analysis of global solar radiation in Nigeria. Archives of Applied Science Research, 2(5): 36-44.

Abdullahi, M. and Singh, S. K. (2014). Global solar radiation evaluation for some selected stations of north Eastern Nigeria. Archives of Physics Research, 5 (2):1-8

Abiodun, C. I. and Anisiji, O. E. (2019). Performance Probability Distribution Function for Modelling Solar Radiation in South Southern Nigeria: A Case Study of Yenagoa. Journal of Energy Research and Reviews, 2(3): 1-8, 2019

Adejumo, A. O., Suleiman, E. A. and Okagbue, H. I. (2017). Exploration of Solar radiation data from three geo-political zones in Nigeria. Data in Brief, 13: 60-68.

Adjekukor J. A. (2019). Estimation of Economic Growth using Export and Import Goods in Nigeria: A Kernel Regression Approach. International Journal of Research and Scientific Innovation (IJRSI) , VI(IV): 305- 308.

Nadaraya, E. A. (1964). On estimating regression. Theory of Probability and its Application, 9:141–142.

Temitope, R. A. (2015). Determination of probability distribution function for modeling global solar radiation: Case study of Ibadan, Nigeria. International Journal of Applied Science and Engineering, 13(3): 233- 245.

Uka, C. O., Osuji, G. A. and Aronu, C. O. (2018). The Performance of the Kernel Regression Model for Assessing the Impact of Money Supply on Industrial Growth in Nigeria. International Journal of Advances in Scientific Research and Engineering (ijasre), 4(7): 32-39.

Appendix

Table 2: Table showing the observed total solar radiation, Predicted solar radiation and error

Month	Total_Solar_radiation	Predict_kmodel	Error
JAN	10.99	10.99002	-2.4E-05
FEB	20.49	20.49002	-2.2E-05
MAR	38.55	38.55001	-8.8E-06
APR	60.07	60.06997	3.21E-05
MAY	69.02	69.01999	7.93E-06
JUN	74.86	74.85996	3.97E-05
JUL	65.15	65.15	-2.8E-06
AUG	56.53	56.53002	-1.8E-05
SEP	55.16	55.16005	-4.9E-05
OCT	72.86	72.85988	0.000118
NOV	44.39	44.39	-2.4E-06
DEC	16.86	16.86007	-7E-05