Volume 5 No.9, September 2015 ISSN 2224-3577
International Journal of Science and Technology

http://www.ejournalofsciences.org

Multiple Linear Regression Model for the Estimation of Sulphur Dioxide in Aba Urban, Nigeria

[1]Agbanyim Akuagwu N., [2] Kalu A. U., [3]Ejike E.N.
[1]Department of Chemistry, Abia State Polytechnic, Aba, Abia State. Nigeria.
[2]Department of Mathematics, Abia State Polytechnic, Aba, Abia State. Nigeria.
[3]Department of Chemistry, Federal University of Technology, Owerri, Imo State, Nigeria.

ABSTRACT

Atmospheric pollution has continued to increase exponentially with increase in domestic and industrial activities in major urban and sub-urban cities of the world. Sulphur dioxide and other oxides of Sulphur are component of atmospheric pollution. It has continued to be a global problem due to its adverse effect on both the environment and on human health. Aba being a growing commercial and industrializing urban city in Nigeria is very much at risk of the negative effect of atmospheric pollution. There is therefore a need to constantly monitor the level of Sulphur dioxide in the Atmosphere to prevent it from reaching a level that will be dangerous for human existence in Aba and its environs. The relationship between Sulphur dioxide and four meteorological variables - Wind Speed, Temperature, Relative Humility and Light Intensity, were investigated in Aba urban, using the multiple linear regression technique. The model was found to be:

$$SO_2 = -670.87 - 10.52WS + 10.84TP + 6.72RM - 0.02LI$$

The analysis of variance (ANOVA) conducted on the model shows that it was statistically significant and so a good model for the estimation of Sulphur Dioxide in the Atmosphere in that area. All the assumptions of the model were checked but none of them, including the assumptions of Normality, Linearity, Homoscedasticity and Independence of residuals were violated.

Keywords: *Multiple Linear Regression Model; Atmospheric pollution, Analysis of variance, Correlation coefficient, Coefficient of determination.*

1. INTRODUCTION

Sulphur dioxide is an invisible toxic gas with a pungent, irritating, and rotten smell. It reacts easily with other substances to form harmful compounds such as Sulphuric acid, Sulphurous Acid and Sulphate particles [1].

The main sources of Sulphur dioxide in the atmosphere are through volcanic eruptions and industrial activities that process materials that contain Sulphur in the generation of electricity from coal, oil or gas that contains Sulphur. Sulphur dioxide is also present in emissions from fuel combustion in motor vehicles and electric generating plants [2].

Sulphur dioxide affects human health when it is breathed in. It irritates the nose, throat, and airways to cause coughing, wheezing, shortness of breath, or a tight feeling around the chest. The effects of Sulphur dioxide are felt very quickly and most people would feel the worst symptoms in 10 to 15 minutes after breathing it in [3]. The most at risk of developing problems when exposed to Sulphur dioxide are people with asthma or similar conditions.[4]. The concentration of Sulphur dioxide in the atmosphere can influence the habitat suitability for plant communities, as well as animal life[5]. Sulphur dioxide emissions are a precursor to acid rain and atmospheric particulates. A 2011 systematic review concluded that exposure to Sulphur dioxide is associated with preterm birth[6].

Aba being a growing commercial and industrializing urban city, with a population of over 3 million people with so many motor vehicles and electric generating sets powered by fuel containing at lot of sulphur is at a high risk of dangerous Sulphur dioxide pollution of the atmosphere. The city also still practices a crude method of removing hairs from livestock skins for food by burning used motor tires which contain a lot of Sulphur.

2. METHODS

Aba is a commercial and industrializing urban city in Abia State, South East of Nigeria. It has a population of about 3 million people and is located on latitude 05^0 18^1N and longitude 07^035^1E. The city has a tropical climate and an average annual temperature of about 28^0C. Two major seasons are experienced: a rainy season between April and October and a dry season between November and March. In recent times it is no longer easy to accurately predict the period of these two seasons due to climatic change resulting from the over increasing depletion of the Earth's Ozone (O_3) layer and the increasing pollution of the Earth's atmosphere by pollutants such as Sulphur dioxide (SO_2), Carbon monoxide (CO), Nitrogen oxides (NO_X), Particulate Matter etc. The crude method of using condemned motor tires to remove hairs from skins of animals for food and the increasing dependence of private electric generators for electricity supply in the city is compounding the problem of Sulphur dioxide pollution of Aba.

Eight different locations within the high activity area of the city were chosen for the study. Within each of the locations five different points were chosen to collect data on sulphur dioxide and four meteorological variables: Wind Speed, Temperature, Relative Humidity and Light Intensity. In all a total of 132

samples were collected in March, May, August and December 2014, to reflect the two seasons in Aba. The pollution attribute and the meteorological variables were measured with automatic samplers in-situ in all the locations. Surlphur dioxide was measured with a Crowncon Gasman automatic gas monitors, model CE89/336/EEC while the four meteorological variables were measured using an Environmental meter, model AE09685.

3. THE MODEL

The Multiple Linear Regression (MLR) technique, being one of the most popular and widely used techniques of multivariable analysis is used to investigate the relationship between the dependent variable, Surphur dioxide ($S0_2$) and 4 meteorological variable, Wind Speed (WS), Temperature (TP), Relative Humidity (RM) and Light Intensity, that constitute the independent variables. The general form of a multiple linear regression model is:

$$Y = b_0 + b_1X_1 + b_2X_2 + \cdots + b_kX_k + e \qquad (1)$$

(Where Y is the dependent variable, $b_{0,\, b_1,\, b_2,\ldots,b_k}$ are linear regression parameters or coefficients that measure the direction and magnitude of the relationship between the associating independent variable and the dependent variable.$X_1, X_2, X_3, \ldots, X_k$ are the independent variables. e is the error term of the linear regression).

Following this general form of the MLR model, the model for the estimation of Sulphur dioxide in Aba urban is:

$$SO_2 = \beta_0 + \beta_1 WS + \beta_2 TP + \beta_3 RH + \beta_4 LI + e \qquad (2)$$

(where SO_2 is Sulphur dioxide, β_0 , β_1 , β_2 , β_3 , β_4 are the regression parameters to be estimated. WS, TP, RH and LI are Wind Speed, Temperature, Relative Humidity and Light Intensity respectively, e is the error term of the regression).

4. ASSUMPTIONS

To complete the specification of our model, the following assumptions are made:

1. The error term e is a real random variable.
2. The random variable e has a zero mean value for each X_i. That is, $E(e_i) = 0$
3. The variance of each e_i is the same for all the X_i values (Homoscedasticity). That is $E(e_i^2) = \sigma_e^2$ (constant).
4. The values of each e_i are normally distributed with mean zero and variance σ_e^2. That is: $e_i \sim N(0, \sigma_e^2)$
5 The values of e_i (corresponding to X_i) are independent from the values of any other e_j (corresponding to Xi). That is None-autocorrelation or Serial independence of the e's. $E(e_i\, e_j) = 0$ for $i \neq j$.
6. Every disturbance term e_i is independent of the explanatory variables Xi. That is: $E(e_i\, X_{1i}) = E(e_i\, X_{2i}) = 0$
7. The explanatory variables X's are not perfectly linearly correlated. That is there is no perfect multicollinear X's
8. The appropriate "aggregation bridge" has been constructed between the aggregate macro-variables used in the function and their individual components (micro-variables).
9. The relationship being studied is identifiable.

The Statistical Package for Social Sciences (SPSS) software was used to analyse the data collected from the 132 points in the 8 locations of the study and the result is presented and discussed below.

5. RESULTS AND DISCUSSION

The result of the data analysis using the SPSS software is here under presented and discussed.

The correlation coefficient (R) is a measure of the relationship between the observed and the predicted values. It lies between -1 and +1, (i.e $-1 \leq R \leq 1$), meaning negative perfect correlation to positive perfect correlation. A value of 0 indicates no correlation [7]. The value of R for our model is on table 1, labeled **Model Summary**. The value of 0.804 indicates a high positive correlation between the observed and the predicted values.

Table 1: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of The Estimate
1	0.804	0.646	0.635	41.58597

The coefficient of Determination (R^2) is the square of the Correlation Coefficient and determines the proportion or percentage of the variance in the dependent variable that is explained by the model. In other words, it determines how well the model fits the data [8]. The information is also found in the output Table 1, labeled **Model Summary**. A value of 0.646 means that 64.6% of the variance in Sulphur dioxide is actually explained by the model. This means that the model is a good fit for the data.

The collinearity diagnostics performed on the variables is presented in the column headed **collinearity Statistics** in Table 2, labeled **Coefficients**. None of the Tolerance values is very low (near 0). Also none of the Variance Inflation Factor (VIF) value is greater than 5. These mean that there is no multicollinearity between the independent variables as assumed.

Table 2: Coefficients

Model	Unstandardized Coefficients(β)	Standardized Coefficients(β_1)	t value	Sig..	Tolerance Values	VIF
(Constant)	-670.868		-2.834	0.0050		
Wind Speed	-10.518	-0.125	-1.491	0.139	0.397	2.516
Temperature	10.835	0.497	2.348	0.020	0.262	4.054
Relative Humidity	6.720	1.100	5.220	0.000	0.263	3.937
Light Intensity	-0.015	-0.122	-1.817	0.072	0.615	1.626

From the column titled **Unstandardized Coefficients** in Table 2, labeled Coefficients, we get the values of the regression coefficients of our regression equation as: b_0 = -670.87; b_1 = -10.52; b_2 = 10.84 ;b_3 = 6.72 ; b_4 = - 0.015. Therefore our regression equation is:

$$SO_2 = -670.87 - 10.52WS + 10.84TP + 6.72RM - 0.02LI$$

With standard error of estimate of 41.59 (from Table 1)

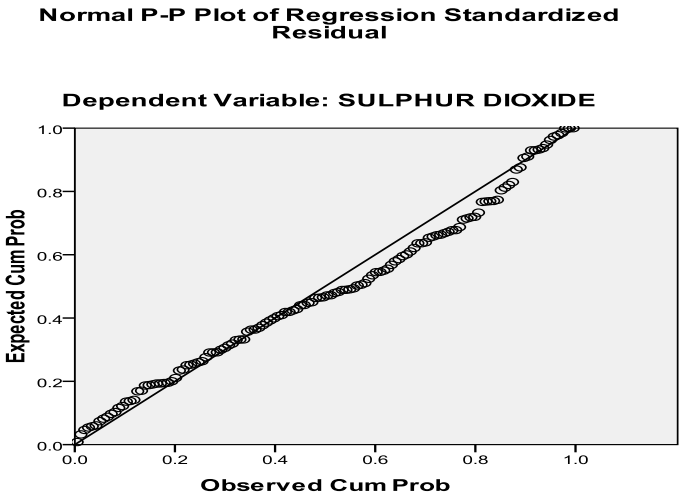
From table 2 also, we find out the relative contribution of the meteorological variables to the variation in SO_2.The contributions from Relative Humidity and Temperature were significant (sig. value of each was less than 0.05). Relative Humidity contributed the most with a Beta value of 1.100 followed by Temperature with a Beta value of 0.497. The contributions from both Wind Speed and Light Intensity were not significant as they have sig. values greater than 0.05.

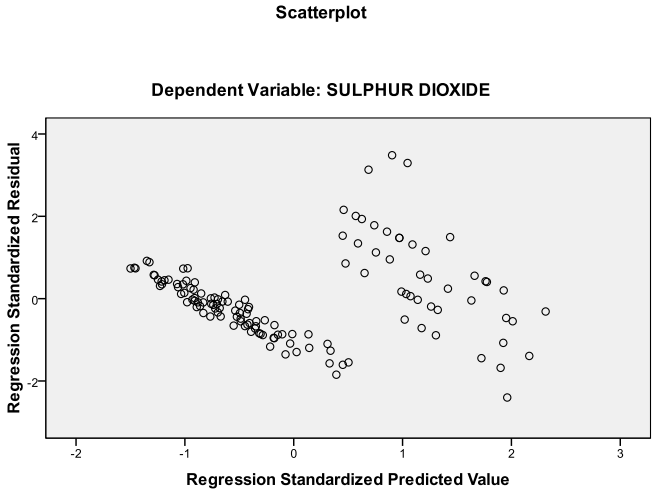
The result of the analysis of variance (ANOVA) performed on the model is presented in Table 3, labeled **ANOVA**. This tests the null hypothesis that multiple R in the population equals O. The significant value of 0.000 indicates that our model reaches statistical significance. This really means P < 0.0005.

Table 3. ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	401274.306	4	100318.577	58.008	0.000
Residual	219632.903	127	1729.393		
Total	620907.210	131			

From the **Normal P-P plot of the standardized Residual** we find out that all the points lie in a reasonably straight diagonal line from bottom left to top right .And from the Scatter Plot we observed that the residuals are roughly rectangularly distributed, with most of the points concentrated at the centre. There is no clear or systematic pattern to the residuals. These two plots show that our assumptions of Normality, Linearity, Homesocedasticity and Indpendence of residuals have not been violated.





6. CONCLUSION

A mathematical model to estimate the level of Sulphur dioxide in Aba urban, South East of Nigeria was found to be $SO_2 = -670.87 - 10.52WS + 10.84TP + 6.72RM - 0.02LI$ .The Correlation Coefficient(R) of 0.804 shows that there is a strong positive relationship between Sulphur dioxide (SO_4) and the meteorological variables- Wind Speed (WS), Temperature (TP),Relative Humidity (RH) and Light Intensity (LI). .

The Coefficient of determination (R^2), shows that 64.6% of the variance in SO_4 is explained by the model. In other words, the model is a very good fit for the data collected and so will be very good in estimating the level of SO_4 in Aba at any given time .

REFERENCES

[1]. Lide, D.R., (2006). CRC Hand book of Chemistry and Physics (87th ed). Boca Ratib, FL: CRC Press. ISBNN 0-8493-0487-3.

[2]. Owen, L. A; Pickering, K. T. (1997) An introduction of Global Environmental Issues. Taylor & Francis. PP 33-ISBN 978-0-203-97400-1

[3]. Liv, D; Jin, H, Tang, C; DU, J (2010) "Sulphur dioxide: a novel gasesous signal in the regulation of cardiovascular functions." Mini-Reviews in Medicinal Chemistry 10 (11): 1039-1045. PMID 20540708.

[4]. Yang R, Yang Y, Dong X, WUX, W. Y. (2014) "Correlation between endogeneous sulphur dioxide and homocysteine in children with pulmonary arterial hypertension associated with congenital heart disease" Zhongua Er Ke Za Zhi (in Chinese) 52 (8): 625-629.

[5]. Hogan, C.M (2010). "Abiotic factor" in Encyclopedia of Earth. Emily Monosson and C. Clevelanc (eds.) National Council for science and the Environment. Washington D.C.

[6]. Shah P.S, Balkhair T, knowledge Synthesis Group on Determinants of Preterm/LBW Birth (2011). "Air pollution and birth outcomes: a systematic review" Environ Int. 37 (2): 498-516.

[7]. Anikender, K; Pramila G. (2011). "Forcasting of airquaility in Delhi using principal component regression technique." Atmospheric pollution Research 2 (2011) 436-444.

[8]. Aron, R. (1984). "Models for estimating current and future sulphur dioxide concentrations in Taipei." Bulletin of Geophysics 25,47-52.