

Modelling of Gully Erosion Site Data in Southeastern Nigeria, Using Poisson and Negative Binomial Regression Models

Ngozi Levi Nwakwasi

Department of Civil Engineering SEET, Federal University of Technology, Owerri, Nigeria

Email address:

nwakwasilevi@gmail.com

To cite this article:

Ngozi Levi Nwakwasi. Modelling of Gully Erosion Site Data in Southeastern Nigeria, Using Poisson and Negative Binomial Regression Models. *Journal of Civil, Construction and Environmental Engineering*. Vol. 3, No. 4, 2018, pp. 111-117. doi: 10.11648/j.jccee.20180304.13

Received: September 3, 2018; **Accepted:** September 21, 2018; **Published:** October 25, 2018

Abstract: The development of gully and other forms of erosion have become the greatest environmental problem facing the people of Southeastern Nigeria. The availability of farm land for agricultural production and construction activities have, been greatly reduced due to soil erosion. This study is set to apply Poisson and negative binomial regression models to identify the major factors that contribute to gully erosion development in Southeastern Nigeria and to ascertain better model suitable for prediction of gully erosion, using secondary data. Maximum likelihood estimation procedure was used to estimate the parameter of the selected model with the number of gully erosion sites as the response variable (Y) and 5-explanatory variable (X's). Also applying the forward selection criteria to the 5-explanatory variables, model 5 is best suitable for forecasting the subject under study. The result of the Poisson regression model showed that there was over dispersion in gully erosion site data since the dispersion parameter (3.677) was greater than 1 hence underestimating the standard error and over estimating the coefficient of the explanatory variable, consequently giving misleading inference. The result of the assessment criteria for Poisson regression model and Negative binomial regression model revealed that the Negative binomial regression model predicts gully erosion soil data better in southeastern Nigeria as considered in this study. Heavy Rainfall (HRF), Extractive Industries (EXI), Excess Farm activities (EFX) are the major contributors to gully erosion site development in southeastern Nigeria, with Heavy Rainfall ranking first. A model suitable for prediction of gully erosion sites in southeastern Nigeria has been developed.

Keywords: Gully Erosion, Rainfall, Soil, Poisson Regression Model, Negative Binomial Regression Model, Akaike Information Criteria (AIC)

1. Introduction

Soil erosion is considered to be a major environmental problem since it seriously threatens natural resources and the environment [1]. Soil loss by runoff is a severe ecological problem occupying 56% of the worldwide area and is accelerated by human induced soil degradation [2]. Soil erosion is a serious environmental, economic and social problem which not only causes severe land degradation and soil productivity loss but also threaten the stability and health of society in general and sustainable development of rural areas in particular[3].

The menace of soil erosion especially gully in no doubt represents a major ecological challenge facing most states in

Nigeria especially Anambra, Imo, Ebonyi, Abia and other states in the humid tropical regions of southern Nigeria [4]. Soil of southern Nigeria has high soil erodibility and is classed as structurally unstable Ofomata [5]. According to [6], the presence of gully site is one of the hazardous features that characterize the Southeastern zone. Gully erosion have become one of the greatest environmental disasters facing many towns and villages in Southeastern Nigeria [7, 8].

A conservative assessment shows the distribution of known gully sites in different stages of development as follows Abia (300), Anambra (700), Ebonyi (250), Enugu (600) and Imo (400). The statistics are not exhaustive enough as new sites are developing during each rainy season due to flooding and torrential rainfall [9, 10]. The average depth of

gullies existing in Ideato North and South L. G. A ranges between 15-35 metres, with a cross sectional area of about 800 metres in some places and covering a distance of about 3km [11]. This paper therefore presents the application of Poisson and negative binomial regression model that predicts gully erosion sites data in southeastern Nigeria to determine the soil erosion parameters that contribute more to gully erosion menace in the region.

A brief review of the causes of gully erosion in Southeastern Nigeria.

From field observation, productive works and reports of other researches carried out within the region, the major causes of gully erosion are as follows:

- (1) Climatic factors: FAO and Okoroafor concluded that gully erosion results from the action of heavy rainfall on surface earth materials under reduced or altered vegetative cover [12, 13]. According to Igwe, the rainfall of southern Nigeria is heavy and aggressive thus generating large volume of runoff that initiate the development of waterways and channels that result to

gully [14].

- (2) Soil Nature and Topography: The Southeastern Nigeria is susceptible to gully erosion due to the nature of the soil, topography and geology [15 - 18]. The Imo/Anambra basin is predominated by the Akwa-orlu cuesta which is an area susceptible to ground surface cracks, landslides, mass movement and tectonic movements during the rainy season that results to all kinds of land degradation and soil erosion predominantly [19, 20].
- (3) Human Factors: According to Egede soil/land has been subjected to intensive pressure from human uses that induced degradation, soil loss and erosion. Such human factors include overgrazing, excessive farm activities, clearing of bushes, extractive industries, road construction, overpopulation, lumbering, residential buildings, development of urban Centre, poor drainage networks etc [21]. Uncontrolled mining operation and soil excavation for various developmental project in the southeast contribute also to gully erosion in the region.

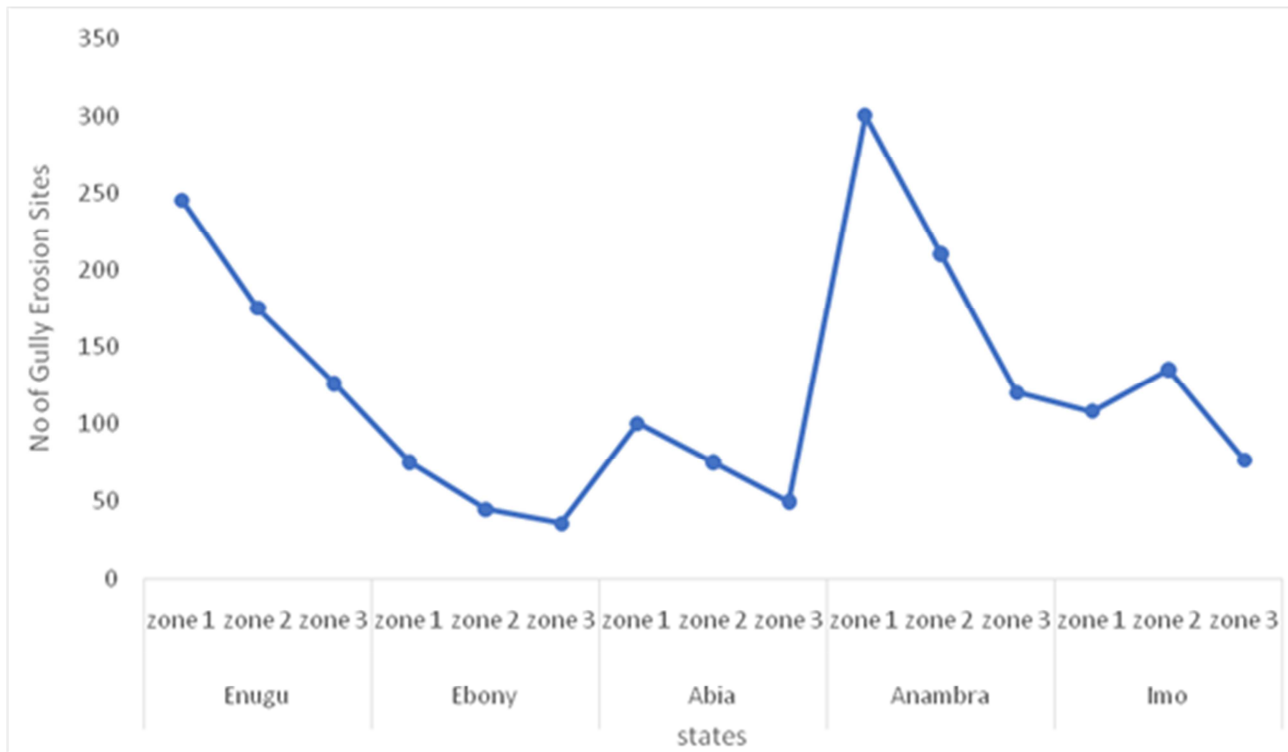


Figure 1. Plot of No of Gully Sites in the Southeastern region.

The objective of this study is to provide a model which will be able to predict the major contributors of gully erosion site in southeastern Nigeria.

2. Methods

2.1. Poisson Regression Model

By definition, Y (dependent variable) follows a Poisson distribution with parameter $\lambda > 0$ iff the probability distribution function is given by

$$P(Y = k) = \frac{\exp(-\lambda)\lambda^k}{k!} \tag{1}$$

For $K=0, 1, 2, \dots$ such that $E(Y) = \lambda$ and $Var(Y) = \lambda$
 For n independent random variables Y_1, Y_2, \dots, Y_n
 $Y_i \sim P(\mu_i)$, and suppose we want to let the mean and the variance depend on the explanatory variables x_i 's

$$g(\mu_i) = \mu_i = x_i' \beta = \eta_i, \tag{2}$$

We can consider a generalized linear model with log link as:

$$\log(\mu_i) = x_i' \beta. \quad (3)$$

Where X_i denotes the vector of explanatory variables and β denotes the vector of regression parameters.

2.2. Model Specification

The model for the Poisson regression is given as:

$$\lambda_i = E(Y_i) = \exp(\alpha + \sum_{j=1}^k x_j' \beta_j), \quad (4)$$

Where α is the intercept, Since the mean is equal to the variance the usual assumption of Homoscedasticity would not be appropriate for a Poisson data.

2.3. Parameter Estimation-Iterative Reweighted Least Squares

Estimation of parameters in Poisson regression relies on maximum likelihood estimation (MLE) method.

The log likelihood function is given as:

$$\log L(\beta) = \sum \{y_i \log(\mu_i) - \mu_i\}, \quad (5)$$

To obtain the maximum likelihood for the parameter β_j we employ the chain rule

$$\frac{\partial l}{\partial \beta_j} = U_j = \sum_{i=1}^N \frac{\partial l_i}{\partial \beta_j} = \sum_{i=1}^N \left(\frac{\partial l_i}{\partial \theta_i} \cdot \frac{\partial \theta_i}{\partial \mu_j} \cdot \frac{\partial \mu_i}{\partial \beta_j} \right) \quad (6)$$

Hence the iterative equation would be written as:

$$b^{(m)} = (X^T W X)^{-1} X^T W \quad (7)$$

Where $w_{ii} = \frac{1}{\text{var}(Y_i)}$ and $z_i = y_i$

2.4. Negative Binomial Regression (NBR) Model

NBR is a popular generalization of Poisson regression because it loosens the highly restrictive assumption that the variance is equal to the mean made by the Poisson model. The traditional negative binomial regression model, commonly known as NB2, is based on the Poisson-gamma mixture distribution. This model is popular because it models the Poisson heterogeneity with a gamma distribution. This is given as:

$$\lambda = E(y_i) = \exp(\alpha + \sum \beta_i X_i + \sum \gamma_i D_i) \quad (8)$$

$$P(Y_i = y_i) = \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!}, y_i = 0, 1, 2, \dots \quad (9)$$

$$Y_i \sim \text{NB}(\lambda = \exp(X^T \beta), \psi) \quad (10)$$

Compounding Poisson distribution above and a gamma distribution would give a Negative Binomial distribution

$$\Pr(Y_i = y_i | \mu_i, \alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1) \Gamma(\alpha^{-1})} \left(\frac{1}{1 + \alpha \mu_i} \right)^{\alpha^{-1}} \left(\frac{\alpha \mu_i}{1 + \alpha \mu_i} \right)^{y_i} \quad (11)$$

Where $E(Y_i) = \mu_i$ while the variance

$$\text{Var}(Y_i) = \mu_i (1 + \alpha^{-1} \mu_i) \text{ where } \alpha = v^{-1} \quad (12)$$

- (1) The regression coefficients are estimated using the method of maximum likelihood.
- (2) The significant regression parameters from (11) will be subjected to a test of hypothesis.
- (3) The AIC would reveal the better model between equation (4) and equation (11) then the rank test would be employed to ascertain the levels of contribution of the various factors under consideration to gully erosion in south eastern Nigeria.

2.5. Model Evaluation

In this research work, we would be applying the techniques below for evaluating the models to also ascertain the better model suitable for modelling gully erosion site data in South-eastern Nigeria.

Deviance: is a goodness of fit statistic for a model that is often used for statistical hypothesis testing and often used to compare two different models. Larger deviance indicates low performance of the model and less deviance indicates better performance of the model.

$$D = 2(l(y, \phi; y) - l(\hat{\mu}, \phi; y)) \quad (13)$$

Where $l(y, \phi; y)$ is the log-likelihood of the full model and $l(\hat{\mu}, \phi; y)$ is the log-likelihood of the current model.

AIC: The Akaike would be used to in selecting the best model that fit our data. The model with the smaller AIC is the best.

$$\text{AIC} = -2(\text{likelihood} - K) \quad (14)$$

$$\text{BIC} = -2\ln(\text{likelihood}) - K \ln(n) \quad (15)$$

Where n is sample size and K is number of predictors;

Pearson Residual: It is used to check for model fit of each variable (explanatory variables). It is the discrepancy between our observed and fitted values for each observation.

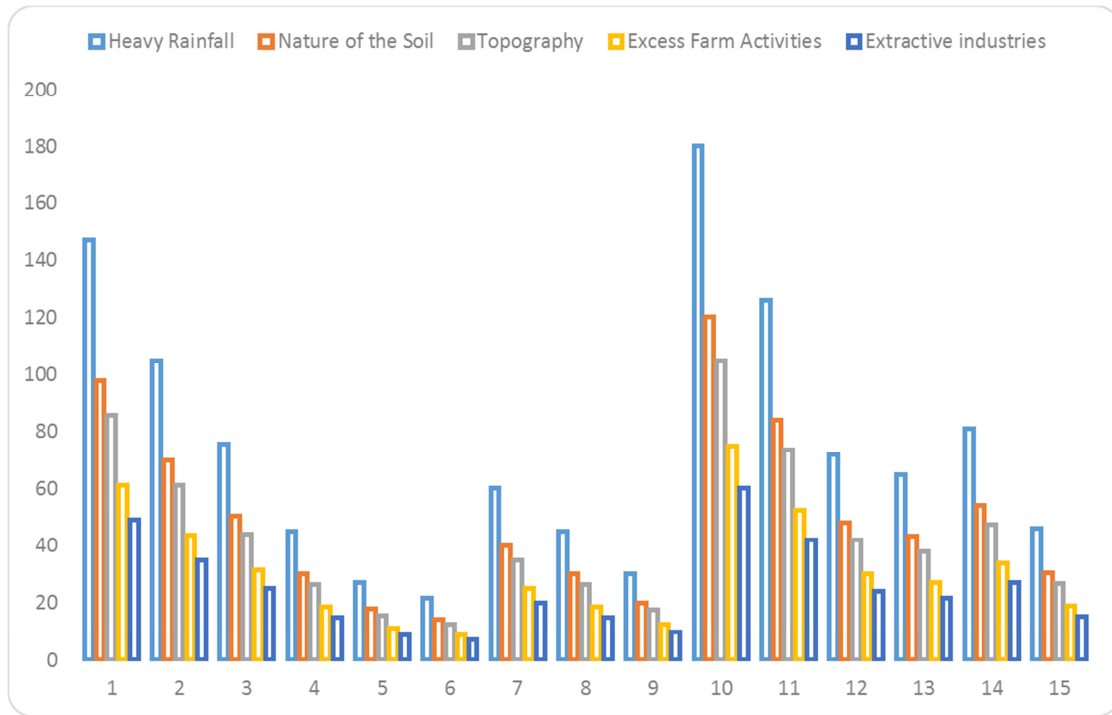


Figure 2. Plot of Contributors to Gully Erosion in Southeastern Nigeria.

3. Results and Discussion

Table 1. Descriptive statistics of the response variable and the explanatory variables.

Descriptive Statistics	N	Minimum	Maximum	Mean	Std. Deviation
No of Gully Sites	15	36	300	125.13	77.016
Heavy Rainfall	15	22	180	75.13	46.183
Nature of the Soil	15	14	120	50.00	30.834
Topography	15	13	105	43.87	26.941
Excess Farm Activities	15	9	75	31.40	19.253
Extractive Industries	15	7	60	25.00	15.432

In modelling the number of gully sites in the Southeastern Nigeria using Owerri, Abia, Enugu as a pivot point for the work, R statistical software version 3.3.3 was used. The Generalized Linear Model (GLM) with Poisson as the fundamental distribution for modelling a count data using the Log link function and the Negative Binomial distribution was latter employed to correct the error of over dispersion in the count data in situation where the result of the Poisson regression model shows over dispersion.

3.1. Modelling Incident of Gully Erosion Data in Southeastern Nigeria

Table 2. The Poisson Regression Models for Gully Erosion in Southeastern Nigeria with their AIC's.

Models	AIC
$\log(\text{mean no of gully}) = \alpha_1 + \beta_1 \text{HRF}$	144.87
$\log(\text{mean no of gully}) = \alpha_2 + \beta_1 \text{HRF} + \beta_2 \text{NOS}$	146.85
$\log(\text{mean no of gully}) = \alpha_3 + \beta_1 \text{HRF} + \beta_2 \text{NOS} + \beta_3 \text{TOP}$	145.07
$\log(\text{mean no of gully}) = \alpha_4 + \beta_1 \text{HRF} + \beta_2 \text{NOS} + \beta_3 \text{TOP} + \beta_4 \text{EFX}$	142.31
$\log(\text{mean no of gully}) = \alpha_5 + \beta_1 \text{HRF} + \beta_2 \text{NOS} + \beta_3 \text{TOP} + \beta_4 \text{EFX} + \beta_5 \text{EXI}$	139.26

HRF=Heavy Rainfall., NOS=Nature of the Soil, TOP=Topography, EFA=Excess farm activities, EXI=Extractive Industries.

From table 2 above, we see that the best model which fit the gully erosion data in Southeastern Nigeria is model 5 because it has the smallest Akaike Information Criterion (AIC). The AIC of model 5 was found to be 139.26 with a null deviance of 622.402 on 14 degree of freedom and a

residual deviance of 33.094 on 9 degree of freedom following the chi-square distribution ($\chi^2_{(9)}$).

The parameters and their estimates, the standard errors, the Chi-squares with their associated probability values are presented in table 3 below

Table 3. The Parameter Estimates of the Selected Poisson Regression Model.

Parameter	Estimate	Std. Error	Z values	Pr(> Z)
(Intercept)	3.84414	0.05694	67.516	<2e-16
Heavy Rainfall	-0.07881	0.10253	-0.769	0.4421
Nature of the Soil	0.04609	0.11816	0.390	0.6965
Topography	-0.18656	0.09308	-2.004	0.0450
Excess farm Activities	0.21402	0.09714	2.203	0.0276
Extractive Industries	0.23675	0.17580	1.347	0.1781

The table 3 presents the parameter estimates of the selected model for the gully erosion in southeastern Nigeria at 5 different states. The Akaike information criterion (AIC) of this model was 142.5 with a null deviances of 622.402 on 14 degree of freedom and a residual deviance of 33.094 on 9 degrees of freedom following the chi-square distribution (χ^2) on one degree of freedom. The dispersion parameter was found to be 3.667 (i.e residual deviance/degree of freedom as seen in table 5) and the Omnibus test flag a P-value equal 0.000 which implies that the model is significant at 5% α -level. However, the assumption of equality of mean and variances in Poisson distribution has been violated since

the dispersion parameter is not approximately equal to 1. The dispersion parameter of the above model is 3.667 which is greater than 1, a clear indication of over dispersion in the gully erosion data. This further implies that the parameters of the stated model have been over-estimated and the corresponding standard errors have been under estimated consequently giving a misleading inference about the regression parameters. To address this issue, Negative Binomial regression was used to modify the model to nullify the effect of over dispersion in the data and the result is shown in table 3 below.

Table 4. Negative Binomial Regression Model Parameter Estimates.

Parameter	Estimate	Std. Error	Z-values	Pr(> Z)
(Intercept)	3.78531	0.08144	46.479	<2e-16
Heavy Rainfall	-0.04861	0.14928	-0.319	0.000474
Nature of the Soil	0.08963	0.16181	0.554	0.04796
Topography	-0.27475	0.14144	-1.943	0.05210
Excess farm Activities	0.19778	0.14862	1.331	0.01832
Extractive Industries	0.23295	0.24240	0.961	0.00336

Interpretation of Coefficients

From table 4 it can be seen that Heavy Rainfall, Nature of the Soil, Excess Farm activities and Extractive Industries were all statistically significant since (p-value < 0.05) while Topography was not statistically significant since (p-value > 0.05).

The variables Heavy Rainfall (HRF), Nature of the Soil (NOS), Excess Farm activities (EFX) and Extractive Industries (EXI) were all statistically significant at 5% α level. This implies that Gully Erosion in the Southeastern region of Nigeria is not as a result of only poor drainage, but from this study we see all the above listed factors

contributing significantly to the increasing numbers of deadly gully erosion as seen or documented within the southeastern region of Nigeria. From this study, it can be established that Heavy Rainfall, Nature of the Soil, Excess Farm activities and Extractive Industries are they major contributors to gully erosion in the southeastern region of Nigeria.

From Table 4 it is observed that the parameter estimates have been reduces and the standard errors have also been increased. The parametric analysis for the comparison between the Poisson and Negative Binomial Regression for goodness of fit of the model is shown in table 5.

Table 5. Assessment Criteria for Poisson and Negative Binomial Regression.

Assessment Parameter	Poisson Regression Model	Negative Binomial Regression Model
Null Deviances	622.402	278.39
Degree of Freedom	14	14
Residual Deviance	33.094	14.38
Degree of Freedom	9	9
Log Likelihood	-137.534	-123.119
Dispersion Parameter	3.6771	1.5977
Akaike's Information Criterion (AIC)	142.5	137.12

From the result presented in table 5, it is clear that the negative binomial regression model is the better model which fits the gully Erosion Data in Southeastern Nigeria since the dispersion parameter has reduced from 3.6771 which was

given by the Poisson model to 1.5977 in the Negative Binomial model.

The Akaike Information Criteria (AIC) of the Poisson Regression model also reduced from 142.5 to 137.12 in the

negative binomial model, we can see more supporting claims from the Null Deviances; Residual Deviance; Log Likelihood all of which also corroborated our findings that the Negative Binomial Model is the better model to modelling the number of Gully Erosion Sites in the Southeastern region of Nigeria.

3.2. The Model for Predicting No of Gully Erosion Sites in Southeastern Region of Nigeria

For negative binomial regression, the model for the number of gully erosion data is obtained as:

$$\log(\text{mean no of gully Erosion Sites}) = 3.78531 - 0.04861(\text{HRF}) + 0.08963(\text{NOS}) - 0.27475(\text{TOP}) + 0.19778(\text{EFX}) + 0.23295(\text{EXI})$$

Table 6. Ranking the Probability Value of the Estimates of Negative Binomial Regression Model.

s/n	Parameter	P-values	Ranks
1	Heavy Rainfall	0.000474	1 st
2	Nature of the Soil	0.04796	4 th
3	Topography	0.0521	5 th
4	Excess farm Activities	0.01832	3 rd
5	Extractive Industries	0.00336	2 nd

From table 6 which ranks the P-values of the better model (Negative binomial Regression) to identify which parameter was more significant also which factor contributed more than the other. Heavy Rainfall was rank in the first position. Implication is that Heavy Rainfall factor is the most statistically significant factor that contributes to high number gully erosion in the southeastern Nigeria followed by Extractive Industries, Excess farm Activities, Nature of the Soil while Topography was not a significant contributor to the subject matter in this study.

4. Conclusion

The extent of soil erosion occurring in the study area is still increasing and is now a major course for concern. From field observation, productive works and reports of other researchers carried out within the region, the major causes of

gully erosion stem from climatic factors, soil nature, topography and human factors. The generalized linear model (GLM) with poison as the fundamental distribution for modelling a count data using the log link function and the negative binomial distribution was later employed to correct the error of over dispersion in the count data in situation where the result of the poisson regression model shows over dispersion. The negative binomial regression model is therefore the better model suitable for modeling gully erosion in the southeastern regions of Nigeria as considered in this study. Heavy Rainfall (HRF), Nature of the Soil (NOS), Excess Farm activities (EFX) and Extractive Industries (EXI) are the major factors that contribute to the high number of gully erosion in southeastern Nigeria. The model which can be used to forecast future number of gully erosion in the region keeping the factors considered in this study constant is shown below:

$$\log(\text{mean no of gully Erosion Sites}) = 3.78531 - 0.04861(\text{HRF}) + 0.08963(\text{NOS}) - 0.27475(\text{TOP}) + 0.19778(\text{EFX}) + 0.23295(\text{EXI})$$

References

[1] J. F. Babatunde, O. A. Olabanji, Y. B. A. Adeleye, I. A. Adegoke and O. Chris, GIS-based estimation of soil erosion rates and identification of critical areas in Anambra sub-basin, Nigeria. Model. Earth Syst. Environ. DOI 10.1007/s40808-016-0218-3. Springer International Publishing Switzerland, 2016.

[2] A. C. Emeribeole and C. J. Iheaturu, Mapping of potential soil erosion risk areas in Imo State using the Revised Universal Soil Loss (RUSLE), Remote Sensing (RS) and Geospatial Information System (GIS) Techniques. International Journal for Research in Emerging Science and Technology, Vol. 2, Issue-10, Oct-2015.

[3] Y. Farhan, Z. Dalal, and F. Ibrahim, Spatial Estimation of Soil Erosion Risk Using RUSLE Approach, RS, and GIS Techniques: A Case Study of Kufranja Watershed, Northern Jordan. Journal of Water Resource and Protection, vol. 5, pp. 1247- 1261, 2013.

[4] Ume, N. C, Enwereuzor, A. I., Egbe, C. A, Ike, M. C. and S. J. Umo (2014). Application of Geographic information system and remote sensing in identifying the impacts of gully eroding in Urualla, Ideato North, Local Government area, Imo state. Nigeria Global Research Journal of Science 3 (3):1-8.

[5] Idowu, O. J and Oluwatosin, G. A. (2008), Hydraulic Properties in relation to morphology of a tropical soil with hardened plinthite under time land use types. Tropical and Sub-tropical Agro Ecosystems 8 (4): 145-155.

[6] Ofomata, G. E. K. (1985). Soil Erosion in Nigeria. The Views of a Geomorphologist University of Nigeria Inaugural Lecture Series No.7 University of Nigeria Press Nsukka, Nigeria.

[7] Adekalu, K. O. Olorunfemi I. A. and J. A. Osunbitan (2007) Grass mulching effect on infiltration, surface runoff and soil loss of three agricultural soils in Nigeria. Bio resource technology, 98(4): 912 – 913.

[8] Okpala A. O (1990). Nigerian population growth and its implications for economic development. Scandinavian Journal of Development Alternatives, 9 (4): 63-77.

[9] Igbokwe, J. I, Akinyede, J. O, Dang, B. Alaga, T, Ono, MN, Nnodu, e2) V.0 and L. O. Anike, (2008), Mapping and Monitoring of the impact of gully erosion in southeastern Nigeria using satellite remote sensing and geographic information system. International Archives of Photogrammetry, Remote Sensing and Spatial Information 38: 113-126.

- [10] Egboka, B. C. E (2004). Distress call and plea to the senate committee for urgent Actions Against Floods, Soil/Gully erosion/landslides Disasters in the southeastern Nigeria, paper presented to senate committee on Environmental; roads/Erosion senate delegation.
- [11] Gunawan G., Dwita S., Herr, S., and W. Sulostiwemi, (2013), Soil Erosion Estimation based on GIS and Remote Sensing for Supporting Integrated Water Resources Conservation Management, *International Journal of Technology*, 2 (3): 147-156.
- [12] FAO (2011). *The State of Food Insecurity in the World, special world report of FAO-UN*, Rome, 2011: 11-14.
- [13] Okoroafor O. O. Akinbile C. O, Adeyemo, A. J., (2017), Soil Erosion in Southeastern Nigeria a Review.
- [14] Igwe, C. A., (2012), Gully Erosion in Southeastern Nigeria: Role of Soil Properties and Environmental Factors; *InTech Open Science*, pp.157-171.
- [15] Ufot, U. O., Iren, O. B., and C. U. Chikere-Njolcu (2016). Effects of and use on Soil physical and chemical properties in Akokwa are of Imo State, Nigeria. *International Journal of Life, Science and Scientific Research*, 2 (3):1-6.
- [16] George, NA, Obot, I. and Akpanetuk, N (2008). Geoelectrical investigation of erosion and flooding using the lithologic compositions of erosion and flood-stricken road in Ukanafun Local Government Area, Akwa Ibom State, Southern Nigeria. *Disaster Advancement.*, 1 (4): 46-51.
- [17] Osadebe, CC, and Akpokodje, E. G (2007). Statistical analysis of variability in properties of soils in gully erosion sites of Agulu-Nanka-Oko area, southeastern Nigeria. *Journal of Mining Geology* 43 (2): 197-202.
- [18] Onu, D. O. (2005). Determinants of farm - level soil conservation and erosion control adoption and utilization behaviours in the ecologically vulnerable areas of Imo State, Nigeria. *Journal of Rural Development* 24 (4): 521 - 543.
- [19] Ezezika, O. O and O. Adetona (2011). Resolving the gully erosion problems in southeastern Nigeria: Innovation through public awareness and connectivity based approaches. *Journal of Soil Science and Environmental management* 2 (10): 286-291.
- [20] Teme S. C and P. O. Youdeowei (2004), Geotechnical investigations for design of foundations for erosion and flood control structures at Unwana Beach, Afikpo, Ebonyi state, Southeastern Nigeria. *Fifth International Conference on Case Histories in Geotechnical Engineering*, April 13-17, 2004. 1: 9. New York, N. Y.
- [21] Egede, E. A (2013). Threats and Mitigation of Soil Erosion and Land Degradation in Southeast Nigeria. *Journal of Environment and Earth Science* 3(13): 95-102.