American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-7, Issue-7, pp-104-112
www.ajer.org
Research Paper Open Access

Multiple Linear Regression Approach To Predicting Noise Pollution Levels and Their Spatial Patterns For The Tarkwa Mining Community Of Ghana

Baffoe[1], P. E. And Duker[2], A. A.
[1](Department of Geomatic Engineering, University of Mines and Technology, Ghana)
[2](Department Geomatic Engineering, Kwame Nkrumah University of Science and Technology, Ghana)
Corresponding Author: Baffoe1, P. E

ABSTRACT: *Predicting and preventing intraurban noise pollution in our communities are very challenging for urban planning, epidemiological studies and environmental management; especially in the developing world. Most existing noise-predicting models are limited in providing changes in noise levels during intraurban development and the corresponding noise pollution. In this study, noise levels were measured at 50 purpose-designed monitoring stations and then a land-use regression model was developed for the intraurban noise prediction applying the multiple linear regression (MLR) approach. The measured and the predicted noise levels were compared; and then further compared with noise estimates from a standard noise model, Lyons empirical model. The results from the developed MLR model did not show any significant differences in the patterns as compared with those of the Lyons Empirical model. The model performance indicators showed a standard deviation of 1.585, high correlation (r) of 0.98,* R^2 *of 0.961 and RMSE of 1.569. The resulting maps showed a heterogeneous distribution of the noise pollution levels in the community. This confirms the usefulness of the method for assessing the spatial pattern of noise pollution in a community. Thus, making it a useful tool for urban planning, epidemiological studies and environmental management.*

KEYWORDS: *Noise Level; Noise Pollution, Noise Models; Land Use Regression Models*

Date of Submission: 28-06-2018 Date of acceptance: 13-07-2018

I. INTRODUCTION

Intraurban noise pollution is prevalent in our cities, especially in the mining communities where many factors contribute to noise pollution and distribution. Increase in urban noise pollution brings about associated health problems including hearing impairment, sleep disturbances, interference with spoken communication, cardiovascular problems, disturbances in mental health, impaired task performance, negative social behaviour and annoyance reactions (Goines and Hagler, 2007). Other studies have also confirmed that noise pollution is a threat to the health and well-being of humans (Benfield, 2012; Goines and Hagler, 2007; Ighoroje *et al*., 2004; Passchier-Vermeer and Passchier, 2000). With the continuous trend of population growth, urbanisation and its associated varied and mobile sources of noise, noise pollution will increase in magnitude and severity. This therefore, calls for more research in this area; however, comprehensive noise exposure assessment techniques are not available. Most noise predicting models, especially the existing Lyons Empirical model which is being applied for noise prediction in the area; are limited in predicting long term noise levels due to the fact that they are limited to only one variable namely, traffic noise. In view of this land-use regression (LUR) models which could be used for a wide range of variables are currently being used; as being used for assessing the exposure to air pollution (Aguilera et al., 2014). This method uses least-squares regression modeling techniques to predict air pollution levels based on the monitoring data accessed at the purpose-designed monitoring stations (PMS) and predictor variables collected mainly through Geographic Information Systems (Aguilera et al., 2014). LUR models are very easy to adopt since they have quality performance in detecting environmental air pollutions in the urban areas; as their empirical structure requires the use of standardized approaches.

In the field of noise exposure, distribution and prediction, application of LUR modelling has been least explored. It was first applied in north-east china, where the technique was used in two different sites; using 101

PML for model development and 101 PML for model validation. The model performance explained 83.2% variability of the noise pollution levels and was successfully used at three different scales (Xie et al., 2011). The second application of the LUR was in three different European cities to explain variability of the intraurban noise pollution of the cities. The model performance was good with adjusted R^2 range of 0.66-0.87 and also 0.70-0.89 in both applications. The short-term noise measurements gave correlation of 0.62-0.78 with noise estimates from the standard noise models as compared with it (Aguilera et al., 2014).

In line with the reported merits of LUR in the literature, the aim of this current study is to develop a generic LUR model for predicting noise pollution levels in urban areas, especially in the mining communities using MLR method. The resulting tools are potential for urban planning, epidemiological studies and environmental management

II. MATERIALS AND METHODS USED

2.1 Study Area

Tarkwa Mining Community (TMC) is an area on the south-western part of Ghana and is within the Tarkwa Nsuaem Municipality. The study area is geographically located between latitudes 5º 17' 00" N and 5º 20' 00" N and longitudes 1º 55' 30" W and 2º 00' 00" W, about 89 km north of Takoradi, the Capital of the Western Region (Mantey and Tagoe, 2012).

It is an old mining town and is well noted for the mining of minerals such as gold and manganese. Goldfields Ghana limited, Anglo-gold Ashanti, and Ghana Manganese Company are some of the large scale mining companies found in the TMC. There are also numerous allied mining companies located there. Several small scale mining activities are also going on in the TMC. Over the past few years TMC has seen infrastructural developments including road constructions, building of health posts, education, industries, banking, hospitality services and private business development (Kumi-Boateng, 2012). Figure 1 shows a map of the Tarkwa Mining Community.

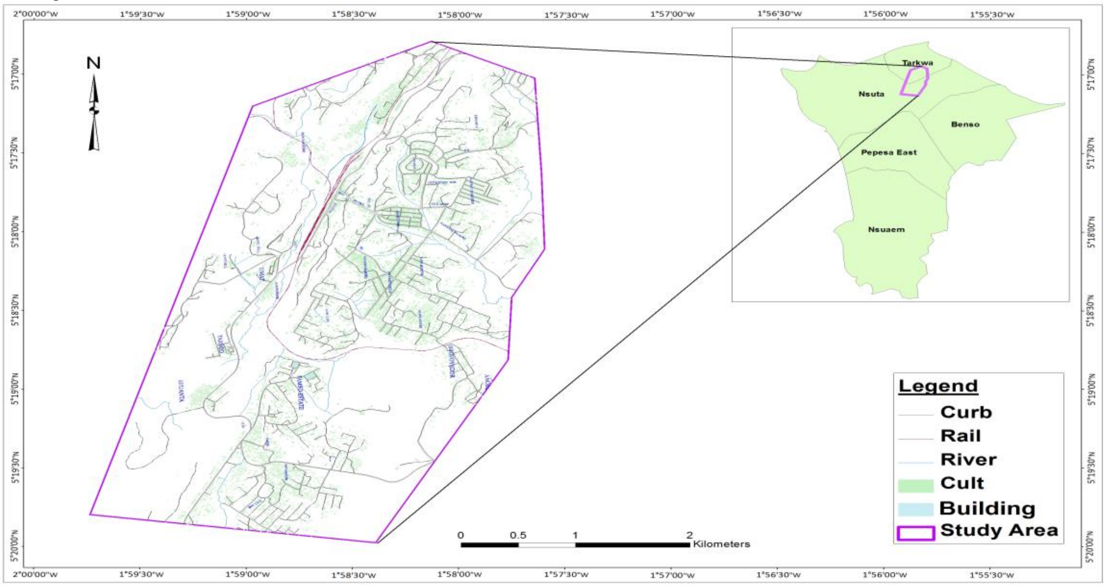

Fig. 1 Tarkwa Mining Community

2.2 Field Data

The geo-spatial locations of the purpose-designed monitoring stations (PMS) in the community were surveyed using Garmin GPS 60CSx handheld Global Positioning System (GPS) of 2 m accuracy. A calibrated Larson Davis's SoundTrack LxT Sound Level Meter was used to measure the noise levels in the study area. The outdoor noise levels were measured from August 2014 to January 2015. The measurements of the PMS were taken at street level and were also determined with the aid of the city digital map of the area.

To avoid noise reflections, the noise-level meter was set on a tripod at about 1.5 m above the ground level and separated from the noise sources by at least 1.5 m. This decision was made in connection with what has been reported and accepted in the literature. For example, Mehdi *et al.*, (2010) used 1.5 m above ground level and 1.22-1.52 m from the source of the noise. The tolerance of the calibrated Larson Davis's SoundTrack LxT trademark device is ±0.6 dBA. A-weighted instantaneous sound pressure levels were recorded three times daily at the selected positions in the study area. The total number of the PMS used for the modeling was 50.

2.3 Predicting Noise Pollution Levels using MLR Approach

The noise predictive model for forecasting noise levels in the study area was developed by following the normal procedures for developing land-use regression model, using the multiple linear regression method. The noise level was used as the dependent variable and the areas of the various land-uses within the study area were defined as independent variables.

Thus, the general equation consisted of five independent variables namely land-use, traffic intensity, road network, distance to the main road, and population density. Due to the heterogeneity nature of the independent variables, they were converted into a homogeneous one that is adaptable in the multiple linear regression equation. This was done using the Analytic Hierarchy Process of the Multi-Criteria Decision Analysis technique; which is used to solve complex multi-criteria decisions. The equations formulated from AHP were solved using matrices.

Application of the Analytic Hierarchy Process

A mathematical expression was developed to provide weights for each criterion as a function of its rank. In computing the vector of criteria weights, a pairwise comparison matrix (A) was formulated for each pair of independent variables. The matrix *A* is m x m matrix, where *m* is the number of evaluation criteria that was considered. Each entry a_{ik} of matrix *A* represented the importance of the *j*th criterion compared to the *k*th criterion. The following conditions were set as defined in AHP procedures: if $a_{ik}>1$, then the *j*th criterion is more important the *k*th criterion and vice versa. If two criteria have the same importance, then the entry a_{jk} is 1. The entries, a_{jk} and a_{kj} satisfy the following constraint in Equation 1:

$$a_{jk} \,.\, a_{kj} = 1 \tag{1}$$

In this study, the importance of each factor criterion over the other was determined and quantified using the scale of pairwise comparison as developed by Saaty (1980) in the AHP.

The comparisons were quantified on a scale of 1 to 9. The quantity 1 represents two factors of equal importance. The quantity 9 also denotes a factor with extreme importance over the other (Saaty, 2008). Based on literature as well as advice from experts' opinion, the judgements for the independent variables were formulated. Each independent variable has five alternatives and five decision criteria. Each of the alternatives was evaluated in terms of the decision criteria and the relative importance (or weight) of each criterion. .

The results are thus represented in normalized pairwise comparison matrices in proceeding Equations 2 and 3.

$$\bar{a}_{jk} = \frac{a_{jk}}{\sum_{l=1}^{m} a_{lk}}. \tag{2}$$

After creating the normalized pairwise comparison matrices for the independent variables, the criteria weight vector, w is built by averaging the entries on each row of A_{norm} that is,

$$w_j = \frac{\sum_{l=1}^{m} \bar{a}_{jl}}{m}. \tag{3}$$

The whole processes were then summarized thus; the judgment tables was represented by 5 x 5 matrices and then squared to obtain an eigenvector. The result was then normalised by summing the eigenvector and dividing each value of the eigenvector by the sum. The weights for the individual factors were obtained after the normalisation process. The process was then repeated a number of times until the weights assigned to each factor were consistent. A consistency ratio of 0.0155012 was achieved which is less than the maximum allowable ratio of 0.10.

Applying the Multiple Linear Regression Approach

The prediction model used to estimate noise pollution levels in the study area was developed using multiple linear regression approach, with matrix notation and analyses using the Statistical Package for the Social Sciences (SPSS) and MATLAB. From the multiple linear regression (MLR) model the dependent variable is related to five independent variables. The general multiple linear regression expression for *k* variables as given in Equation 4 as:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + e_i, \quad i = 1, 2, \ldots, n. \tag{4}$$

In Equation 4, y is the dependent variable which is the noise pollution levels in the study area; β_1, β_2 and β_k are coefficients of the regression equation and β_0 is the intercept that is the value of y when all xs are zero.

In order to solve for the regression coefficient in the MLR model, as in Equation (4) , the method of least squares is used. These coefficients (see Equaton 4), illustrate the unrelated contributions of each independent variable towards predicting the dependent variable. It is important to note that, the computations used in finding the regression coefficients (βi, i = 1, 2, 3, …, k), residual sum of square (SSE), regression sum of squares (SSR), etc is complex therefore, the multiple regression model in terms of observations were written using matrix notation. This is because using matrix allows for a more compact framework in terms of vectors representing the observations, levels of regressor variables, regression coefficients, and random errors. Therefore, Equation 4 was represented in a compact for in Equation 5 as:

$$Y = \mathrm{A} + X\beta + \epsilon \tag{5}$$

Where, Y is Noise Pollution levels, X represents the independent variables; ϵ is Residuals; and A is the value of Y when all Xs are zero. After formulating the matrix equations, Matrix Commands in Excel were used to solve the equations and the results compared with that from SPSS software. Finally using MATLAB, the values for the coefficient of regression **β,** were calculated and the prediction equation, Equation 6 denoted by:

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + \acute{\varepsilon} \tag{6}$$

The least square estimator of the coefficients of the regression equation (β) is given by the Equation 7:

$$\beta = (X^TX)^{-1}(X^TY) \tag{7}$$

Since X is not usually a squared matrix, it is multiplied by the transpose of X, that is X^TX and the inverse of (X^TX) calculated. Hence the estimator β calculated thus using Equation 7.The data was populated using the MATLAB software and the predicted noise levels as projected for the future were then used to develop the spatial distribution of the estimated noise levels.

Applying the Lyons Empirical Model

In order to confirm the developed MLR noise prediction model, the Lyons Empirical prediction model was compared to it. The Lyons Empirical Model is a mathematical model (Sandeep *et al*., 2005) developed and applied in an area of similar conditions as that of the TMC. The long term estimates of noise pollution levels by this model were achieved by using Equation 8:

L (Road Noise) = 10log**V** - 15log**D** + 30 log**S** + 10log [tanh (1.19*10-3) ***VD/S**] + 29 (8)

Where,
L (Road Noise) = Predicted noise level (dBA),
V = Volume of traffic per hour (vehicles/hour),
S = Average vehicle speed (km/hour),
D = Distance from Centre line of road.

Although the Lyons Empirical Model is a mathematical model, in this research it was considered as a "GIS only" model; since all the computations and developments were applied in the GIS environment using the Raster Calculator in ArcGIS. The mapped domain was in grid cells and the Map Algebra was used to do the calculations. Formulating the noise map using this model, the average volume of vehicle per hour counted was 520 vehicles/hour. The average speed determined was 45 km/h and a distance map of 7 meters cell size was used. All the roads in the TMC are single carriage therefore the 7 meters cell size was used since their width is about 7 meters. The data were input into the Lyons Empirical Model and the raster calculator was used to calculate the noise map.

III. RESULTS AND DISCUSSION

3.1 Results of the Measured Noise Levels

The results of the measured noise levels are presented in Table 1. These are presented with locations of the PMS indicated 'Sites' along the major road (as explained earlier) and the computed values of noise level descriptors for the various locations in the study area.

Table 1 Average Noise Descriptors at the Monitoring Locations

Sites	L_{Aeq} dBA	L_{10} dBA	L_{90} dBA	TNI dBA	L_{NP} dBA
1	65	68	58	68	75
2	78	81	73	75	86
3	84	88	74	100	98
4	84	86	75	89	95
5	75	77	67	77	85
6	86	88	79	85	95
7	79	81	72	78	88
8	88	92	76	110	104
9	85	91	80	94	96
10	86	91	79	97	98
11	89	92	81	95	100
12	90	94	82	100	102
13	91	96	85	99	102
14	98	100	90	100	108
15	96	99	86	108	109
16	94	96	81	111	109
17	83	85	75	85	93
18	81	85	74	88	92
19	84	87	72	102	99
20	85	90	71	117	104
21	75	81	69	87	87
22	76	79	68	82	87
23	74	78	65	87	87
24	77	79	72	70	84
25	79	84	72	90	91
26	74	78	65	87	87
27	73	79	65	91	87
28	86	88	76	94	98
29	88	91	82	88	97
30	84	87	78	84	93
31	89	93	77	111	105
32	87	90	79	93	98
33	89	92	78	104	103
34	90	94	77	115	107
35	95	97	81	115	111
36	98	101	88	110	111
37	97	100	79	133	118
38	87	91	78	100	100
39	86	92	78	104	100
40	89	94	75	121	108
41	93	97	84	106	106
42	95	99	82	120	112
43	94	100	85	115	109
45	96	101	88	110	109
46	92	98	85	107	105
47	88	92	82	92	98
48	80	85	71	97	94
49	76	78	71	69	83
50	68	73	63	73	78

Where LAeq is A-weighted equivalent sound pressure level; L10 and L90 are the exceedence percentiles, average sound level, LD, the day-night average sound level, LDN, the noise pollution level, LNP and the traffic noise index, TNI.

3.2 Results from the Prediction Model

The results of the MLR model developed for predicting the noise levels of the study area are presented. Regression analysis allows modelling, examining, and exploring of spatial relationships. This helps to better understand the factors behind observed spatial patterns, and to predict outcomes based on that understanding. The resulting correlation coefficient between the variables, summary coefficients of the multiple linear regression and others from the SPSS statistical software are presented in Tables 2 and 3. The results of the summary of the parameters of the model as ran in SPSS version 2010.

Table 2 Model Summary

Mathematical Model	Performance Indicators						
	RMSE	Mean	Standard Deviation	Maximum	Minimum	R^2	R
MLR	1.569	2.462	1.585	3.718	-3.756	0.961	0.98

RME = root mean square error, R^2 = coefficient of determination, R = correlation coefficient

Table 3 Coefficients of the Regression Equation

	Coefficients	*Standard Error*	*Lower 95%*	*Upper 95%*
Intercept	68.088	1.0811	65.907	70.268
POP	112.602	4.9542	102.610	122.593
TRAFFIC	-124.815	5.7975	-136.506	-113.123
ROADNET	36.570	2.1050	32.324	40.815
LANDUSE	14.146	2.1729	9.763	18.528
DIST	45.503	2.9734	39.506	51.500

The noise prediction model developed from the hybrid approach is given in Equation 9.

Noise Levels = 112.602*(POP) – 124.815*(TRAFFIC) + 36.570*(ROADNET) + 14.146*(LANDUSE) + 45.503*(DIST) + 68.088 (9)

The deterministic component that is in the form of a straight line provides the predicted (mean/expected) response for a given predictor variable values. The residual terms represent the difference between the predicted value and the observed value. They are assumed to be independently and identically distributed normally with zero mean and variance; and account for natural variability as well as measurement error.

The predicted noise levels from the MLR model was compared with that of those measured from the field and the results are present in Figure 2. Figure 2 shows the trend of predicted and measured values of noise level during day-night in the study.

.

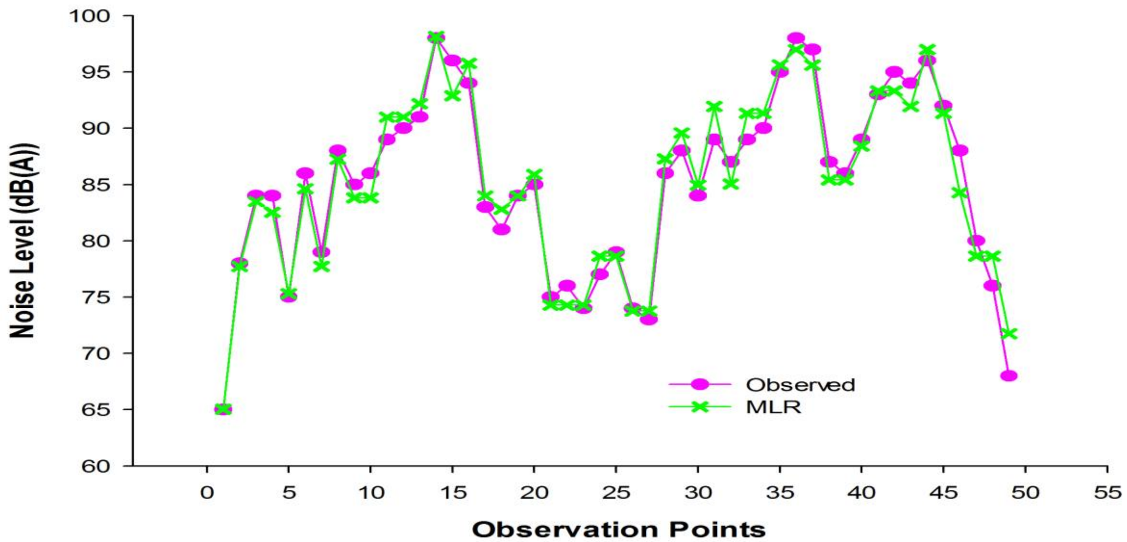

Fig. 2 Trend of Predicted and Measured of Noise Levels for the Study Area

The developed MLR model was also run in Matlab and the results were used to plot the spatial distribution of the noise levels of the area. The spatial distribution of the estimated noise pollution levels from the monitoring stations in the study area forecasted are presented in Figure 3 demonstrating that GIS could be used for noise mapping.

The standard model, Lyon's Empirical model, which was used to confirm the developed MLR model, was also used to map the TMC and the result presented in Figure 4. Since this model is a highway noise prediction model, whose principle is based on the fact that noise is produced by traffic and then reduced by distance before reaching the listener, all the major roads were used. The predicted noise levels were between 65.94 dB (A) being minimum level and 110.42 dB (A) being the maximum level.

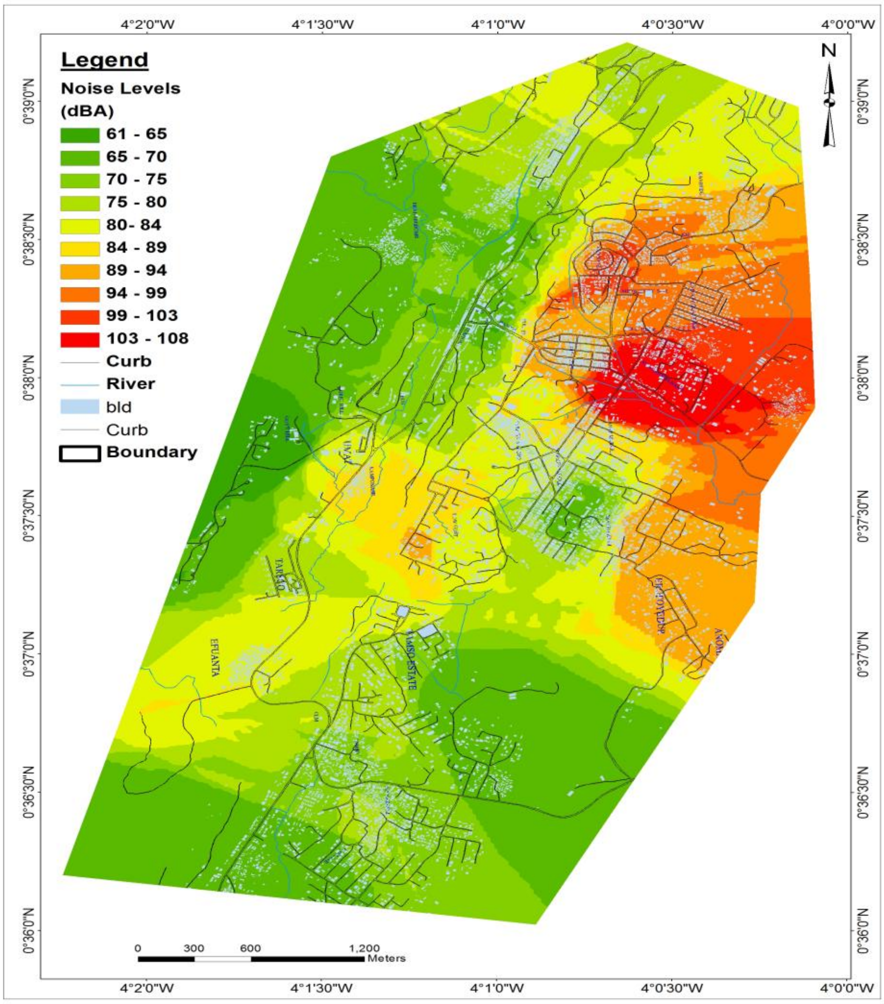

Fig. 3 Noise Map for the Predicted Noise Levels

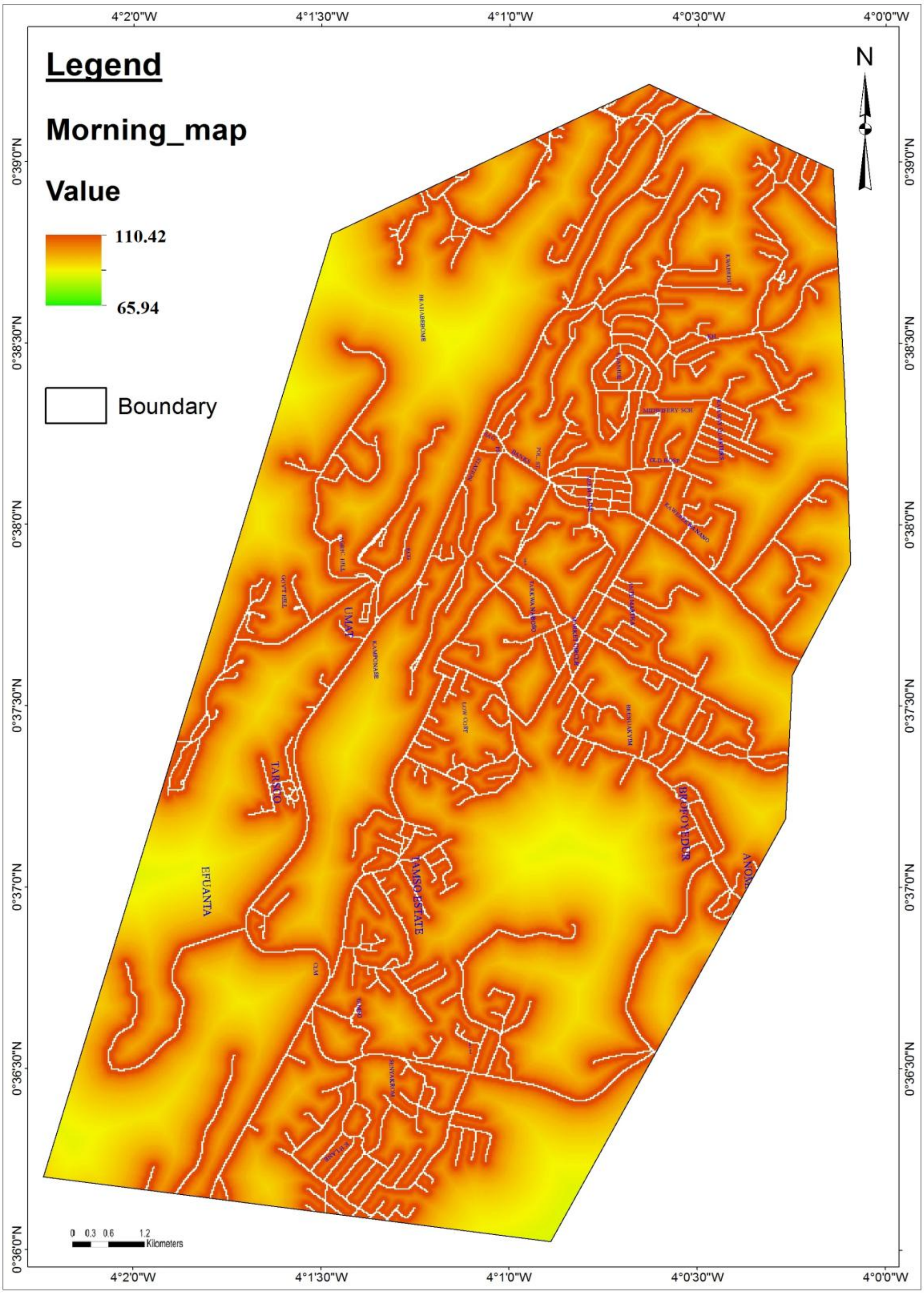

Fig. 4 Noise Map Created from Lyons' Empirical Noise Prediction Model

IV. CONCLUSIONS

A noise prediction model has been designed and implemented, the first time of its kind in the West African sub-region, using MLR method, which could be replicated, in other areas. The statistical indicators used to verify the viability of the developed MLR model showed a RMSE of 1.569, standard deviation of 1.585, correlation coefficient (R) of 0.98 and R^2 of 0.961. These statistical findings have revealed that the developed LUR based on the MLR could be used for noise prediction of an area with an accuracy of 98%. This assertion is based on the R of 0.98 obtained in this study. The use of R as a model assessor is in line with the study of Xie et al., (2011) and Aguilera et al., (2015) who also used the same approach for the model adequacy. Furthermore, results obtained from the MLR model is in consonance with the standard noise model (Lyons Empirical model). This was noticed from the visual observation from the noise map produced in the standard as shown in Figure 3 and 4. Hence, the developed MLR model had also demonstrated its use in mapping intraurban noise pollution levels in relation to urban land-use as changes occur. This is useful for urban planning and environmental noise management. It could also be applied predict intraurban noise changes with time and epidemiological studies as well as decision-making tool. It was observed from the study that the more the PMS, the better the model performance. Therefore, the model performs better for large scale area like the study area and vice versa.

REFERENCES

[1]. Aguilera, I., Foraster, M., Basagaña, X., Corradi, E., Deltell5, A., Morelli, X., Phuleria, H. C., Ragettli, M. S., Rivera, M., Thomasson, A., Rémy Slama, R., and Künzli, N., (2015), "Application of Land-use Regression Modelling to Assess the Spatial Distribution of Road Traffic Noise in Three European Cities.", Journal of Expo Science Environ Epidemiol. Vol. 25 (1): 97-105.

[2]. Goines, L. and Hagler, L. (2007), Noise Pollution: A Modern Plaque, Southern Medical Journal, Vol. 100, Pp. 287-294

[3]. Ighoroje, A. D. A, Marchie, C., and Nwobodo, E. D. (2004), Noise-Induced Hearing Impairment as an Occupational Risk Factor among Nigerian Traders, Nigerian Journal of Physiological Sciences 19(1-2): 14-19, Physiological Society of Nigeria 2004.

[4]. Kumi-Boateng, B. (2012), "A Spatio-Temporal Based Estimation of Land-Use Cover Change and Sequestered Carbon in the Tarkwa Mining Area of Ghana", PhD Thesis, University of Mine and Technology, Tarkwa, 2012, pp. 163.

[5]. Passchier-Vermeer, W., and Passchier, W. F., (2000), Noise Exposure and Public Health, Environmental Health Perspective, Vol. 108, p. 123-131.

[6]. Saaty, T. L. (1980), The *Analytical Hierarchy Process*, McGraw-Hill, pp. 23-157.

[7]. Xie, D., Liu, Y., and Chen, J., (2011), "*Mapping Urban Environmental Noise: A Land Use Regression Method"*, Environmental, Science and Technology, Vol. 45, 7358–7364.

Baffoe1, P. E"Multiple Linear Regression Approach To Predicting Noise Pollution Levels And Their Spatial Patterns For The Tarkwa Mining Community Of Ghana."American Journal of Engineering Research (AJER), vol. 7, no. 07, 2018, pp. 104-112.