JOURNAL OF ENVIRONMENTAL DESIGN (JED)

A Journal of the Faculty of Environmental Studies, University of Uyo, Uyo, Nigeria
Vol. 6, No. 1, May, 2011

EDITORIAL STYLE GUIDE FOR AUTHORS

Please submit your manuscript double spaced in MS Word. Provide tables and figures in a separate file (or files) in MS Excel, MS Word, or EPS format. Maps should be supplied in EPS format.

Because manuscripts will undergo a blind review, submit two title pages, the first showing the title of the manuscript, author name, title, affiliation, telephone number, e-mail address, and the date of the manuscript. The second title page should contain only the title of the paper.

Third-person style is always preferred. If appropriate, authors may make limited use of first-person singular, but a single author should not refer to himself or herself as "we."

Biography. The manuscript should include, on a separate page or the "first" title page described above, a sentence listing each author's name, title, and affiliation.

Abstract. Include a two-paragraph abstract *not exceeding 150 words* and place it on the first page of the text. In the first paragraph of the abstract, describe the issue(s) or question(s) the paper addresses. In the second paragraph, state the major findings or conclusions.

Keywords. To help users reference the JED published research, keywords are included with journal articles. Please suggest three keywords for your manuscript.

Abbreviations. The definition of an abbreviation or acronym is given the first time it appears; afterward, only the abbreviation is used. However, an abbreviation that is defined in the abstract should also be defined in the article. An abbreviation that appears only once in an article should be deleted and the full wording used.

If an abbreviation is first defined in the text, the abbreviation alone can then be used in subsequent footnotes or tables; however, if the abbreviation is first defined in a footnote or table, the abbreviation should be defined again when it first appears in the following text.

Text Headings. Headings are not numbered and are placed to the left. First-level headings are bold; second-level headings are italic; and third-level headings are italic with a period that leads directly into text.

Examples:

First-levelheading
Second-levelheading
Third-level heading. The text continues…

Tables and Figures. Use arabic numerals to number tables and gures consecutively in separate series in order of appearance. Include a brief descriptive title at the top of each. Tables and gures should be in separate electronic les, not integrated into the text. The text must contain a reference to each table not integrated into the text. The text must contain a reference to each table or figure.

Equations. Make sure that all symbols in equations are clear and that all equations (except those in footnotes) are numbered. Single-letter variables should be italicized. Multiple-letter variables and abbreviations (e.g., AGE) and functions (e.g., exp, min, In) should not be italicized; neither should numbers, parentheses, or math operations. Vectors and matrices should be in bold (not italicized).

References. The manuscript must include complete and accurate citations of all materials referenced in the manuscript that are not of your original authorship. Please double-check your references to ensure that names and dates are accurate, that Web pages are still active, and that there are no discrepancies between the text and the reference list the APA style is recommended.

TABLE OF CONTENT

APPLICATION OF GIS-BASED ORDINARY LEAST SQUARES AND GEOGRAPHICALLY WEIGHTED REGRESSION TO LAND USE CHANGE MODELING: A CASE OF LAGOS, NIGERIA

Onuwa Okwuashi[a1], Etim Eyo[b] & Aniekan Eyo[c]
(Onuwaokwashi@yahoo.com, etim.eyo@newcastle.ac.uk, &
ani_eyo@yahoo.com)
Department of Geoinformatics & Surveying, Faculty of Environmental Studies, University of Uyo, Nigeria
Department of Geomatics University of Newcastle, Newcastle, United Kingdom

ABSTRACT

This research explored the processes of land use change in Lagos, Nigeria using the geographic information systems based ordinary least squares and geographically weighted regression models, based on three epochs: 19631978, 1978-1984, and 1984-2000. ArcGIS and MATLAB software were used for the modelling. Only two of the five tested statistical assumptions of linear regression modelling were met. Multicollinearity and autocorrelation criteria were met while normality, linearity, and homoscedasticity criteria were not met. Nonetheless, the geographically weighted regression predicted maps for Lagos indicated substantial agreement with the reference data based on the computed Kappa statistic. The results of the modelling on one hand elucidated the tremen[rl0]us explanatory properties of linear regression models; on another hand the results showed that the nature of land use data makes their use for urban change modelling inconsistent with traditional statistical assumptions; therefore the trade-off for deriving highly accurate results from linear regression models is the inevitable violation of traditional statistical assumptions.

INTRODUCTION

Changing land use globally is a topical issue of discussion. Urbanisation may be the benchmark for measuring economic growth and development, however in the case of sprawling cities of developing countries it has been accompanied by poverty, unemployment, environmental degradation, decaying infrastructure, and uncontrollable growth of informal settlements (Angotti, 1993). Urban sprawl in Lagos has put profound pressure on housing, infrastructure, and the environment (Braimoh & Onishi, 2007), and is generally viewed by most Nigerians as an intractable problem (Abiodun, 1974; Gandy, 2006). This research therefore explores the use of the Geographic Information Systems (GIS) based Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) models for modelling the processes of urban change in Lagos, Nigeria.

Conventional land use change modelling is implemented in the ArcGIS environment using the Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) models. The GWR (Fotheringham et al., 2002) is the local equivalent of

$$y_i = \beta_0(u_i, v_i) + \sum_q \beta_q(u_i, v_i)x_{iq} + \varepsilon_i$$

(2)

where β_0 denotes the intercept, β_q represents the slope coefficients for the variable q, x_{iq} denotes the value for a qth variable for i number of observations, ε_i represents the error parameter, and

the global OLS. OLS and GWR are both linear regression models. The global OLS model can be expressed mathematically as,

$$y_i = \beta_0 + \sum_p \beta_p x_{ip} + \varepsilon_i$$

(1)

where y_i are the dependent variables, β_0 denotes the intercept, β_p are the slope coefficients for the p variable, x_{ip} denotes the value for a p th variable for i number of observations, and ε_i represents the error parameter.

Unlike the GWR model, OLS assumes that β_p is stationary or homoscedastic. This is a major difference between the OLS and the GWR (Fotheringham et al., 2002; Mennis, 2006). The GWR model can be expressed mathematically as,

(u_i, v_i) stands for the coordinates of the ith location for an i observations.

GWR takes the effect of spatial dependency into consideration and assumes that β_q is non-stationary or heteroscedastic. The non-stationarity of β_q means that the solutions of β_q vary across the globe for the same values of x_{iq}. The solution of β_q is affected by the locations (u_i, v_i) where x_{iq} were actually observed (Fotheringham et al., 2002).

In spite of the merits of the GWR model over the OLS model, it is still subject to fundamental statistical assumptions (just like the OLS), that there is: (i) a normal distribution; (ii) a linear relationship between the dependent and independent variables; (iii) no multicollinearity between the independent variables; (iv) no spatial autocorrelation; and (v) homoscedasticity (Leung et al., 2000; Fotheringham et al., 2002; Wheeler & Tiefelsdorf, 2005; Farber & Páez, 2007).

Methodology

The study area for this experiment is Lagos, Nigeria (Figure 1). Lagos is a littoral environment, has a relatively flat terrain, an area of about 2910km^2, and lies between latitudes 6° 26′ and 6° 50′ N, and between longitudes 3° 09′ and 3° 46′ E (Braimoh & Onishi, 2007). Substantial land use change has occurred in Lagos between 1963 and 2000 (Figure 2).

Twelve salient land use drivers were selected for Lagos. The selected land use drivers influencing land use change in Lagos are: water, residential structures, industrial and commercial centres, major roads, railway, Lagos Island, international airport, international seaport, University of Lagos, Lagos State University, income potential, and population potential.

The land use data of Lagos consist of two remotely sensed Landsat Thematic Mapper images, acquired in 1984 and 2000 respectively; and two analogue base maps acquired in 1963 and 1978. The analogue base maps were sourced from the Lagos State Ministry of Lands.

The Landsat images acquired in 1984 and 2000 were classified with the *k*-means algorithm using MATLAB software. The analogue base maps acquired in 1963 and 1978 were scanned and digitised using ArcGIS. The analogue and remote sensing data were geo-referenced to ensure both data were in the same coordinate system. Digitising analogue maps reduces their accuracy. It is therefore important to enhance the accuracies of the analogue data to ensure they approximate those of the satellite data. The enhancement of the analogue base maps was done in MATLAB, by first obtaining *n*-classifications of the satellite data and overlaying the resulting classified satellite data with the digitised base maps. A digital editing procedure was then used to remove errors from the digitised maps.

The land use independent variables (Table 1) are grouped into two categories: (i) proximity variables, and (ii) weighted variables. The proximity variables were extracted with the GIS while the weighted variables were extracted in MATLAB.

The proximity variables were extracted by calculating the Euclidean distances from the proximity variables to all cells. For the weighted variables, income potential was estimated by ranking major towns in Lagos, using a ranking r_i of a production/service centre c_i. The ranking was normalized to weight w_i using the formula (Braimoh & Onishi, 2007),

$$w_i = \frac{r_i}{\sum_{c_i=1}^{n} r_i} \quad (3)$$

Therefore the weighted inverse distance formula for calculating the income potential i_S of a location S is given as (Braimoh & Onishi, 2007),

$$i_s = \frac{\sum_{x=1}^{n} \frac{1}{|S - S_x|^2} w_i}{\sum_{x=1}^{n} \frac{1}{|S - S_x|^2}} . \quad (4)$$

where $x = 1,2,...,n$ are settlements in the study area.

The final ranking was based on government documents, interviews, and personal knowledge of the author. The population estimate was obtained using the formula (Braimoh & Onishi, 2007),

$$P_p = \frac{P_r}{e^{r(Y_r - Yp)}} \quad (5)$$

where, r = intercensal population growth rate, P_p = previous population, P_r = present population, Y_r = recent year, and Y_p = previous year.

An inverse distance weighting formula for calculating an unknown population potential p_S, at a given location S, is given as (Braimoh & Onishi, 2007),

$$p_s = \frac{\sum_{x=1}^{n} \frac{1}{|S - S_x|^2} P(S_x)}{\sum_{x=1}^{n} \frac{1}{|S - S_x|^2}} \quad (6)$$

where $P(S_x)$ is the population of a settlement located at S; $x = 1,2,...,n$ are settlements in the study area.

The original land use maps (in 1963, 1978, 1984, and 2000) had a cell size of 100m x 100m before they were gridded to increase their cell size to 500m x 500m (see the overlaid gridded maps in Figure 2). The essence of increasing the cell size of the maps is to reduce the matrix size of the maps for modelling convenience.

The GIS-based GWR and OLS models do not need to train the data before the data is used for prediction. The observed data in 1963 were used to predict the 1978 land use; 1978 data were used to predict the 1984 land use; and the 1984 data were used to predict the 2000 land use. The present year is furnished by the independent variables, while the target year is furnished by the dependent variables. Empirical measurements are only done on the present year. The dependent variables are represented with discrete variables. The present year and target year maps were overlaid to determine the changed regions between the present and target years. The results of the overlay are three categories of land use of the target year: (i) undeveloped region; (ii) change region; and (iii) developed region. The resulting overlaid maps form the data for the dependent variable. The dependent variable was represented as follows: undeveloped region=1; changed region=2; and developed region=3.

The original values of the independent variables were scaled to [0, 1] using the transformation formula (Gong, 1996; Li & Yeh, 2002).

$$x_i' = (x_i - \min)/(\max - \min) \quad (7)$$

where x_i' is the scaled land use variable; $\min$ is the lowest value in the land use vector; $\max$ is the highest value in the land use vector; and x_i represents the land use variables.

This scaling technique is effective in ensuring that all the independent variables are equally weighted (Gong, 1996; Li & Yeh, 2002).

Equations 8-10 and equations 11-13 are the OLS and GWR equations for periods 1963-1978, 1978-1984, and 1984-2000 respectively;

$LUC\_1963-1978_i$, $LUC\_1978-1984_i$, and $LUC\_1984-2000_i$ represent Land Use Change (LUC) from 1963-1978, 1978-1984, and 1984-2000 respectively:

$$LUC\_1963\text{-}1978_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \beta_{10} x_{10} + e_i \quad (8)$$

$$LUC\_1978\text{-}1984_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \beta_{10} x_{10} + e_i \quad (9)$$

$$LUC\_1984\text{-}2000_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \beta_{10} x_{10} + \beta_{11} x_{11} + \beta_{12} x_{12} + e_i \quad (10)$$

$$LUC\_1963\text{-}1978_i = \beta_0 + \beta_1 x_1(u_1,v_1) + \beta_2 x_2(u_2,v_2) + \beta_3 x_3(u_3,v_3) + \beta_4 x_4(u_4,v_4) + \beta_5 x_5(u_5,v_5) + \beta_6 x_6(u_6,v_6) + \beta_7 x_7(u_7,v_7) + \beta_8 x_8(u_8,v_8) + \beta_9 x_9(u_9,v_9) + \beta_{10} x_{10}(u_{10},v_{10}) + e_i \quad (11)$$

$$LUC\_1978\text{-}1984_i = \beta_0 + \beta_1 x_1(u_1,v_1) + \beta_2 x_2(u_2,v_2) + \beta_3 x_3(u_3,v_3) + \beta_4 x_4(u_4,v_4) + \beta_5 x_5(u_5,v_5) + \beta_6 x_6(u_6,v_6) + \beta_7 x_7(u_7,v_7) + \beta_8 x_8(u_8,v_8) + \beta_9 x_9(u_9,v_9) + \beta_{10} x_{10}(u_{10},v_{10}) + e_i \quad (12)$$

$$LUC\_1984\text{-}2000_i = \beta_0 + \beta_1 x_1(u_1,v_1) + \beta_2 x_2(u_2,v_2) + \beta_3 x_3(u_3,v_3) + \beta_4 x_4(u_4,v_4) + \beta_5 x_5(u_5,v_5) + \beta_6 x_6(u_6,v_6) + \beta_7 x_7(u_7,v_7) + \beta_8 x_8(u_8,v_8) + \beta_9 x_9(u_9,v_9) + \beta_{10} x_{10}(u_{10},v_{10}) + \beta_{11} x_{11}(u_{11},v_{11}) + \beta_{12} x_{12}(u_{12},v_{12}) + e_i\,. \quad (13)$$

The meaning of the explanatory variables (x_1, x_2, x_3, x_4 x_5, x_6, x_7, x_8, x_9 x_{10}, x_{11}, x_{12}) are given in Table 1. From equations 8-13, β_0 is the intercept; $\beta_1,...,\beta_{12}$ are the coefficients of the independent variables; u, v are the horizontal coordinates; and e_i is the error term. The data were prepared in MATLAB, and thereafter imported into ArcGIS for modelling. The dependent variables must be shapefiles consisting of discrete variables. The attribute tables of the dependent variables (to be predicted) must contain: unique or primary keys, all the attributes of the independent variables (explanatory variables), and discrete variables of the dependent variables. The dependent variables cannot be represented with continuous variables because the resulting map-overlay consists of only three land use types: undeveloped, change, and developed.

3. Results

The OLS model was used to explore the significance of each explanatory variable (see Tables 2-4). The calculated multiple R^2 values for periods 1963-1978, 1978-1984, and 1984-2000 were 0.318959, 0.360174, and 0.381849 respectively; while the calculated adjusted R^2 values for periods 1963-1978, 1978-1984, and 1984-2000 were: 0.317984, 0.359258, and 0.380787 respectively. The hypothesis for assessing the significance of each explanatory variable can be stated as follows :

H_0: the coefficients are zero at the 95% CL

H_1: the coefficients are not zero (reject H_0 if p-value $<5\%$)

Let us now assess OLS model with respect to some fundamental traditional statistical assumptions. We can now assess how the GIS-based OLS modelling results meet the following traditional statistical criteria: multicollinearity; linearity; normality; spatial autocorrelation; and homoscedasticity or stationarity.

For the multicollinearity test, the Variance Inflation Factor (VIF) (see the last columns of Tables 2-4) assesses the effect of multicollinearity in the spatial model. Explanatory variables with VIF >7.5 are considered redundant, and should be excluded from the model. The calculated VIF values for all the explanatory variables in the three experiments (1963-1978, 1978-1984, and 1984-2000) were <7.5 (see the last columns of Tables 2-4). This implies that all the variables were important in the prediction, and should be retained in the model.

For the linearity test, the scatter plot in Figure 3a depicts a typical relationship between a dependent and an independent variable in linear regression modelling. Based on empirical land use change data, Figure 3b is the plot of a dependent variable (1963-1978) against an independent variable (distance to water 1963-1978). Figure 3b shows that land use change data do not conform to a typical linear regression relationship between a dependent and an independent variable shown in Figure 3a. Figure 3b shows that the relationship between the land use dependent variable (y-axis) and the independent variable (x-axis) is nonlinear as expected. Using discrete variables to represent the dependent variable will definitely produce a nonlinear relationship between the dependent and explanatory variables, as in the case of Figure 3b. Therefore the land use change data do not meet the linearity condition expected of all linear regression modelling.

For normality test, the Jarque-Bera Statistic test (Table 5) is used to test whether the model residuals are normal:

H_0: the residuals are normally distributed at the 95% CL
H_1: the residuals are not normally distributed (reject H_0 if p-value <0.05)

All the calculated *p*-values for 1963-1978, 1978-1984, and 1984-2000 were <0.05; which indicates that the residuals deviate from the normal distribution as expected.

For autocorrelation test, the Moran's I Tool was used, Moran I values close to +1 indicate positive spatial autocorrelation; values close to -1 indicate negative spatial autocorrelation; while values close to zero indicate that the model residuals are random. The Moran's I tests (see Table 6) for 1963-1978, 1978-1984, and 1984-2000 yielded values close to zero. These results indicate that the model residuals were not spatially autocorrelated.

All the three results for periods 1963-1978, 1978-1984, and 1984-2000 indicate significant non-stationarity. This implies that the model is not homoscedastic and therefore violates traditional statistical requirements that expect OLS models to be homoscedastic. In ArcGIS, the GWR model becomes a veritable option for the prediction when the OLS results indicate significant non-stationarity.

The R^2 was used to assess the goodness-of-fit for predicting land use change in the three periods (1963-1978, 1978-1984, and 1984-2000). The resulting high R^2 and R^2 adjusted values (Table 8) show that the GWR predictions for the three periods (1963-1978, 1978-1984, and 1984-2000) were highly accurate.

This result is an indication that the GWR model was able to mitigate the effects of non-stationarity and autocorrelation that usually compromise the goodness-of-fit of the dependent variable as in the case of the OLS model.

The GWR model predicted maps for 1963-1978, 1978-1984, and 1984-2000 and their computed Kappa statistic are given in Figure 4 and Table 9 respectively.

It is desired that the standard residuals (that is under and over prediction) be randomly distributed. The standard residuals for the three periods shown in Figure 5 indicate that the residuals were random. This is an indication that no key independent variable was omitted from the model (Thapa & Murayama, 2009). High local R^2 values (Figure 5) indicate areas on the map where the GWR model predicted well while low local R^2 values indicate areas on the map that were less well predicted by the GWR model. Cond denotes condition number. Cond assesses local collinearity in the model. The results from cells with Cond values greater than 30 may not be reliable. All the predicted Cond maps for periods 1963-1978, 1978-1984, and 1984-2000 (Figure 6) yielded Cond values below 30. This indicates that the results from this prediction can be trusted.

4. Conclusion

The result of this research showed that only some of the linear regression assumptions were met. Often times, traditional statistics has questioned the validity of any modelling result that is inconsistent with traditional statistical assumptions governing linear regression models. Nonetheless, spatial statistics tends to undermine the relevance of traditional statistics in spatial modelling. In recent times, researchers have proposed the adoption of nonparametric models as a way-out of the controversy between traditional statistics and spatial statistics. The result of this work will assist urban planners in Lagos to: (i) understand how each land use driver affects the process of urban change and, (ii) forecast urban change between epochs.

REFERENCES

Abiodun, J. O. (1974). Urban growth and problems in metropolitan Lagos. *Urban Studies, 11* (33), 341-347.

Angotti, T. (1993). *Metropolis 2000: Planning, poverty and politics*. London, England: Routledge.

Braimoh, A. K. & Onishi, T. (2007). Spatial determinants of urban land use change in Lagos, Nigeria. *Land Use Policy, 24*(2), 502-515.

Farber, S. & Páez, A. (2007). A systematic investigation of cross-validation in GWR model estimation: Empirical Analysis and Monte Carlo simulations. *Journal of Geographical Systems, 9*(4), 371-396.

Fotheringham, A. S., Brunsdon, C. & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester, England: John Wiley & Sons Ltd.

Gandy, M. (2006). Planning, anti-planning and the infrastructure crisis facing metropolitan Lagos. *Urban Studies, 43*(2), 371–396.

Gong, P. (1996). Integrated analysis of spatial data from multiple sources: Using evidential reasoning and artificial neural network techniques for geological mapping. *Photogrammetric Engineering & Remote Sensing, 62*(5), 513–523.

Leung, Y., Mei, C., & Zhang, W. (2000). Testing for spatial autocorrelation among the residuals of the geographically weighted regression. *Environment and Planning A, 32*(5), 871-890.

Li, X. & Yeh, A. G. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science, 16*(4), 323-343.

Mennis, J. (2006). Mapping the results of geographically weighted regression. *The Cartographic Journal, 43*(2), 171-179.

Thapa, R. & Murayama, Y. (2009). Land use change factors in Kathmandu Valley: A GWR approach. In B. G. Lees & S. W. Laffan (Eds.), *10th International Conference on Geocomputation* (pp. 255-260). UNSW, Sydney, November 30th – December 2nd.

Wheeler, D. & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems, 7*, 161–187.

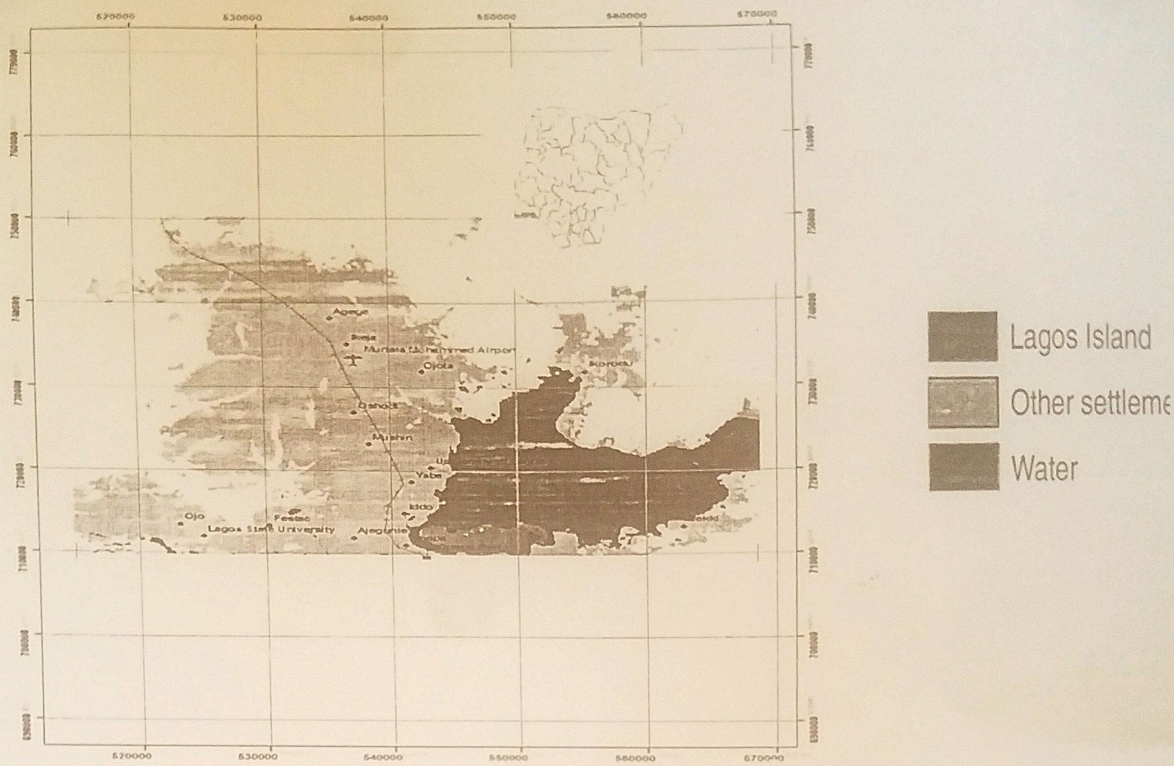


Table 1 *Extracted land use variables*

	Land use variables
Proximity variables	x_1 : distance to water
”	x_2 : distance to residential structures
”	x_3 : distance to industrial and commercial centres
”	x_4 : distance to major roads
”	x_5 : distance to railway
”	x_6 : distance to Lagos Island
”	x_7 : distance to international airport (1984-2000 only)
”	x_8 : distance to international seaport
”	x_9 : distance to University of Lagos
”	x_{10} : distance to Lagos State University (1984-2000 only)
Weighted variables	x_{11} : income potential
”	x_{12} : population potential

Figure 1 *Lagos in relation to Nigeria*

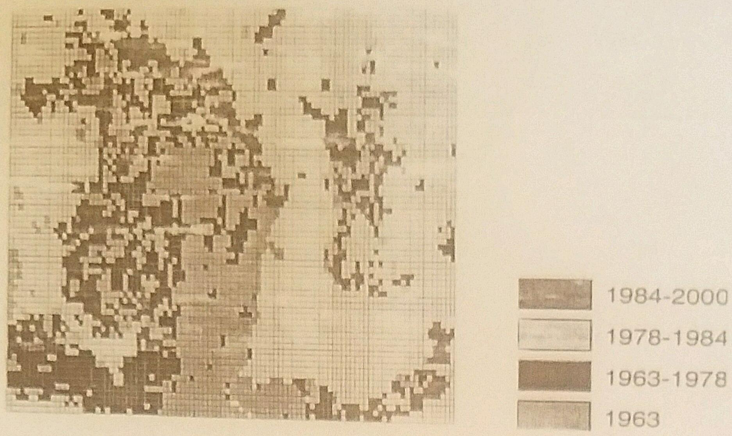


Figure 2 *Gridded land use map of Lagos between 1963 and 2000*

Table 2 *Statistical results for assessing the significance of each independent variable in the model for 1963-1978 (*significant at p<0.05 or t>1.96)*

Variable	Coefficient	Std Error	t-statistic	P-value	VIF [1]
Intercept	0.847101	0.018673	45.363946	0.000000*	- - - - -
Distance to water	-0.117348	0.055944	2.097593	0.035964*	1.509747
Distance to residential	-1.124156	0.050378	22.314590	0.000000*	1.159984
Distance to industrial and commercial	-0.806965	0.045379	17.782810	0.000000*	1.342798
Distance to major roads	-0.421015	0.047971	8.776518	0.000000*	1.271059
Distance to railway	-0.716223	0.042507	16.849342	0.000000*	1.194308
Distance to Lagos Island	-0.397699	0.042050	9.457734	0.000000*	1.325239
Distance to international seaport	-0.174235	0.039471	4.414274	0.000013*	1.254065
Distance to University of Lagos	-0.062781	0.041584	1.509759	0.131166	1.274922
Income potential	0.520350	0.043490	11.964778	0.000000*	1.141347
Population potential	0.230385	0.035664	6.459927	0.000000*	1.141712

Table 3 *Statistical results for assessing the significance of each independent variable in the model for 1978-1984 (*significant at p<0.05 or t>1.96)*

Variable	Coefficient	Std Error	t-statistic	P-value	VIF [1]
Intercept	1.092893	0.022225	49.173422	0.000000*	- - - - - - -
Distance to water	-0.208187	0.065870	3.160574	0.001597*	1.440754
Distance to residential	-1.919315	0.062447	30.735343	0.000000*	1.244771
Distance to industrial and commercial	-1.010004	0.059138	17.078679	0.000000*	1.410719
Distance to major roads	-0.927202	0.074672	12.417024	0.000000*	1.086095
Distance to railway	-0.655001	0.052848	12.393995	0.000000*	1.270764
Distance to Lagos Island	-0.308435	0.050105	6.155783	0.000000*	1.321559
Distance to international seaport	-0.140088	0.047237	2.965640	0.003041*	1.236370

Distance to University of Lagos	0.137918	0.049860	2.766108	0.005689*	1.261713
Income potential	0.746732	0.054971	13.584038	0.000000*	1.178389
Population potential	0.339526	0.057042	5.952223	0.000000*	1.182373

Table 4 *Statistical results for assessing the significance of each independent variable in the model for 1984-2000 (*significant at p<0.05 or t>1.96)*

Variable	Coefficient	Std Error	t-statistic	P-value	VIF [1]
Intercept	-2.290807	0.023219	98.658942	0.000000*	- - - - - - -
Distance to water	-0.222466	0.075164	2.959756	0.003099*	1.176389
Distance to residential	-2.606189	0.077746	33.522009	0.000000*	1.104094
Distance to industrial and commercial	-2.113546	0.092167	22.931596	0.000000*	1.137527
Distance to major roads	-1.486506	0.106371	13.974726	0.000000*	1.063784
Distance to railway	-0.910023	0.058224	15.629605	0.000000*	1.050858
Distance to Lagos Island	-0.339675	0.056222	6.041727	0.000000*	1.093061
Distance to international airport	-0.035236	0.054862	0.642261	0.520725	1.102546
Distance to international seaport	-0.189215	0.062306	3.036886	0.002412*	1.093976
Distance to University of Lagos	0.147534	0.060228	2.449619	0.014312*	1.104240
Distance to Lagos State University	-0.719737	0.058858	12.228258	0.000000*	1.052826
Income potential	0.721885	0.063363	11.392814	0.000000*	1.154016
Population potential	0.414852	0.048891	8.485296	0.000000*	1.153289

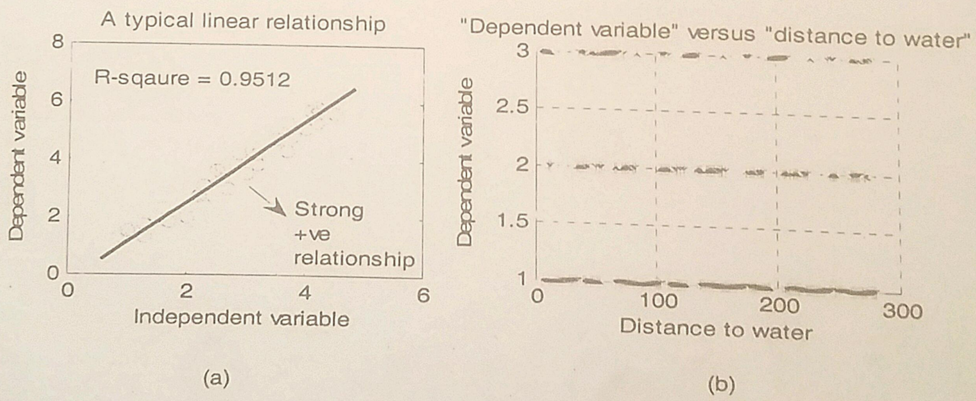


Figure 3 *(a) A typical linear regression relationship between a dependent and an independent variable; (b) dependent variable plotted against distance to water*

Table 5 *Jarque-Bera Statistic test for 1963-1978, 1978-1984, and 1984-2000 (*significant at p<0.05)*

Periods	Jarque-Bera Statistic	Degrees of freedom	*P*-value
1963-1978	706.116841	10	0.000000*
1978-1984	392.961529	10	0.000000*

1984-2000	318.762167	12	0.000000*

Table 6 *OLS modelling: Spatial autocorrelation test for 1963-1978, 1978-1984, and 1984-2000 using ArcGIS Moran's I tool*

Periods	Moran's I index for OLS
1963-1978	0.081073
1978-1984	0.073943
1984-2000	0.069293

For non-stationarity test, the Koenker (BP) Statistic test (Table 8) was used:

H_0: the model is stationary at the 95% CL
H_1: the model is non-stationary (reject H_0 if p-value <0.05)

Table 7 *Koenker (BP) Statistic for 1963-1978, 1978-1984, and 1984-2000 (*significant at p<0.05)*

Periods	Koener (BP) Statistic	Degrees of freedom	*P*-value
1963-1978	1237.735416	10	0.000000*
1978-1984	532.003124	10	0.000000*
1984-2000	74.834390	12	0.000000*

Table 8 *GWR modelling: calculated R^2 values for 1963-1978, 1978-1984, and 1984-2000*

Periods	R^2	R^2 Adjusted
1963-1978	0.843159	0.792102
1978-1984	0.846021	0.805692
1984-2000	0.853782	0.799722

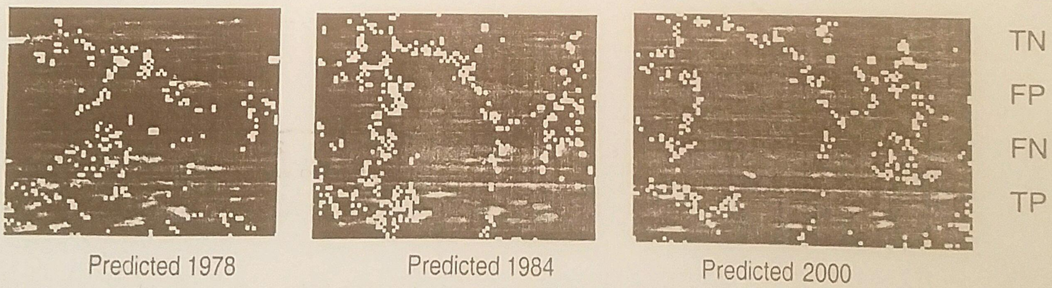


Figure 4 *Predicted GWR maps for 1963-1978, 1978-1984, and 1984-2000 (TN=True negative; FP=False Positive; FN=False Negative; TP=True Positive)*

Table 9 *Calculated Kappa sta*[illegible] *1984, and 1984-2000*

Periods	Kappa statistic
1963-1978	0.8858

1978-1984	0.8366
1984-2000	0.8812

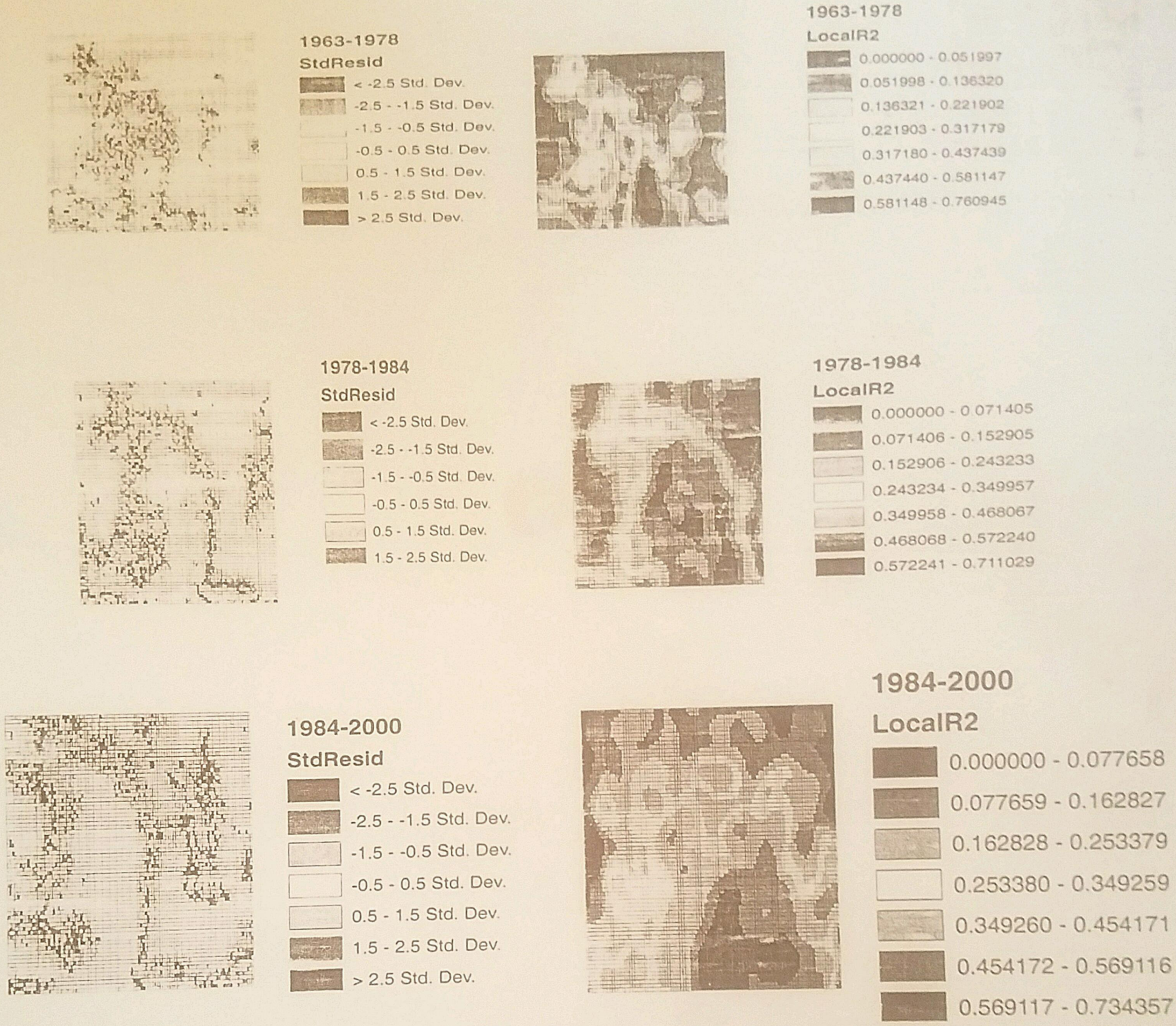
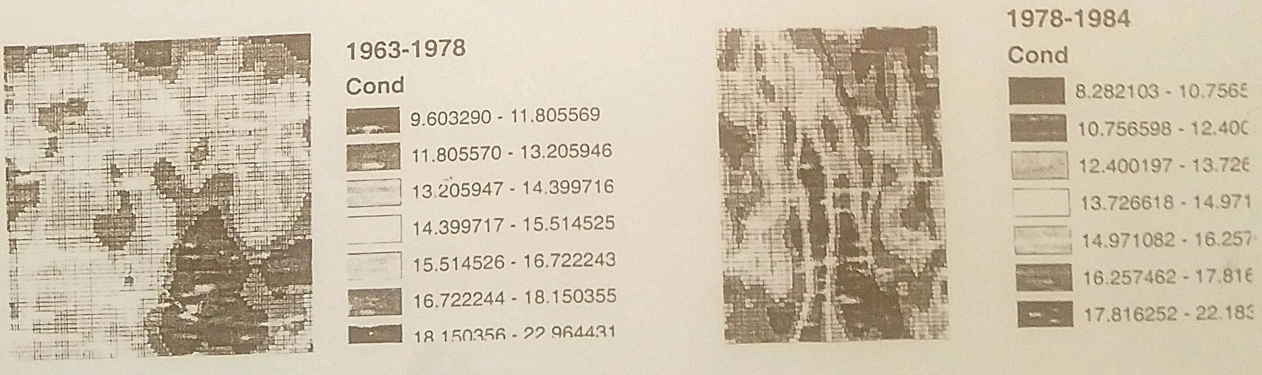


Figure 5 *Estimated GWR standard residuals and local R² for 1963-1978, 1978-1984, and 1984-2000*



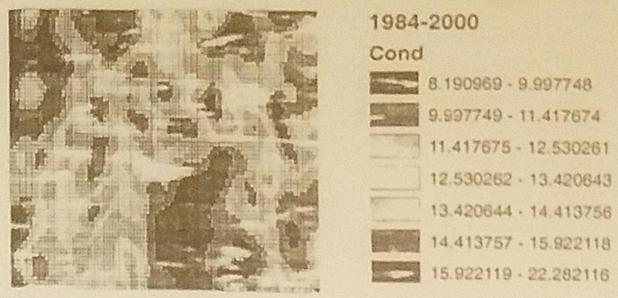


Figure 6 *Estimated GWR model condition numbers for 1963-1978, 1978-1984, and 1984-2000*