Applied Geography 108 (2019) 47–63

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Geographically weighted logistic regression approach to explore the spatial variability in travel behaviour and built environment interactions: Accounting simultaneously for demographic and socioeconomic characteristics

Felix Ndidi Nkeki*, Monday Ohi Asikhia

Department of Geography and Regional Planning, University of Benin, P.M.B. 1154, Benin City, Nigeria

ARTICLE INFO

Keywords:
Geographically weighted logistic regression
Travel behaviour
Built environment
Entropy index
Land use mix
Mode choice

ABSTRACT

The relationship between built environment and travel behaviour has been a topical subject of academic debate over the last two decades. This has given rise to a plethora of empirical literature in this area of study. Ultimately, these studies were conducted using statistical models that fail to explain spatial non-stationarity of the processes in their dataset. To improve understanding concerning built environment and travel behaviour interactions, local model against global model is suggested. The aim of this study is to analyze the spatial variation in travel behaviour and built environment interactions using Geographically Weighted Logistic Regression (GWLR) and at the same time accounting for the individual attributes of commuters (e.g., demographic and socioeconomic characteristics). Based on valid responses from 1028 survey points carried out in Benin metropolitan region, a GWLR of travel mode choice was estimated. The result shows that unlike global statistics, local model revealed a significant spatial variation in the association between travel mode choice and the factor scores of demographic and socioeconomic variables across neighbourhoods. GWLR model also revealed the occurrence of spatial mismatch between demographic and socio-economic characteristics, and this created a dichotomy by demarcating the neighbourhoods into two levels of influence. The result further showed that built environment variables are weak predictors of mode choice in the region. Local model proved to be most suitable for exploring this relationship since it accounted for local variation which is often lost when using global models.

1. Introduction

The interaction between urban built environment and travel behaviour has provoked global responsiveness over the last two decades. Though studies on such association has been linked to the work of Mitchell and Rapkin (1954), many contemporary scholars and associated practitioners still show keen interest in the study of this relationship perhaps, owing to the growing importance and effectiveness of using the resultant information for policy reasons and achieving desirable transport objectives including the issue of excessive carbon emission from car exhaust and over-reliance on fossil fuel. Indeed, the non-spatial and spatial changes that take place in the urban environment have created diverse problems that need to be managed to achieve sustainable development. The need to understand these problems has given rise to a plethora of literature which cut across various regions and diverse disciplines.

Transport planners and other applied geographers have dedicated a chunk of their time and energy to study the relationship between travel behaviour and series of explanatory variables which are classified into individual and built environment attributes. Ultimately, these studies were carried out using statistical models that fail to account for spatial non-stationarity of the processes in their dataset (Rahul & Verma, 2017; Sun, Yan, & Zhang, 2017; Zhang, Yao, & Liu, 2017; Zwerts, Allaert, Janssens, Wets, & Witlox, 2010). From the geographers' perspective, understanding local variability in a relationship is key to remedying spatial disparities. However, global models of travel behaviour are hinged on the choice of individual decision-maker who is seen as the actor assessing the benefits and cost of their travel choices (Pike & Lubell, 2016) and the personal and environmental factors that influence such behaviour. These global models often return aggregated result which is generalised for an entire region of interest. Clearly, the parameters in such models assume fixed relationships between the built environment, individual characteristics and travel behaviour across space. Hence, spatial heterogeneity or disparity in relationships is lost.

To address the limitation of global models of travel behaviour and built

* Corresponding author.
E-mail address: nkekifndidi@gmail.com (F.N. Nkeki).

https://doi.org/10.1016/j.apgeog.2019.05.008
Received 10 October 2018; Received in revised form 21 May 2019; Accepted 24 May 2019
Available online 30 May 2019

environment interactions, Geographically Weighted Logistic Regression (GWLR) approach (credited to Brunsdon, Fotheringham, & Charlton, 1996) is suggested. This is because the parameters of the model are allowed to vary over a geographic area of interest and thereby highlighting the spatially varying relationships between variables. Yet, no earlier studies have accounted for the spatial dimension of this relationship or explore the local variability using spatial logistic model. In other words, there is a dearth of literature on this topic that adopted a location-specific approach while using a logistic modelling framework for data analysis. The underlying idea is to understand the spatial (i.e. the built environment) and non-spatial (i.e. the individual attributes) factors that influences travel behaviour and how such influence varies in space. The need to improve understanding on spatially varied relationships over a geographic area of interest is increasingly gaining momentum among geographers and transport planners (Brunsdon et al., 1996; Du & Mulley, 2012; Wang & Chen, 2017). The reason is credited to the usefulness of the derived information which serves as a valuable tool for strategic planning and policy, such as resource allocation and projection, transport infrastructural development and adjustment, land use planning and restructuring, etc.

In this study, the spatial variation in travel behaviour and built environment interactions were explored using GWLR and simultaneously account for the individual attributes of commuters which are composed of their demographic and socioeconomic characteristics. The GWLR model was computed for Benin Metropolitan Region (BMR) to explore the intrinsic relationships that may exist between the explanatory variables and travel mode choice, focusing more on the significant spatial varying relationships at the neighbourhood-level which global models could not achieve. BMR is characterised by varying travel behaviour which is inherent in the travel mode choices of people across neighbourhoods. For example, some neighbourhoods are mostly car dependent and yet some are essentially bus dependent. These different choices in travel mode utilisation may be influenced by factors that are equally heterogeneous across neighbourhood space. It is expected that when global models are used for this kind of analysis, neighbourhood-level relationships may be masked. Since it is assumed that the neighbourhoods may present distinct pattern of relationship when results are disaggregated. By employing GWLR to analyse the relationship between the travel behaviour of the commuters in BMR and their various individual and built environment characteristics, it is expected that each neighbourhood may present distinct pattern of relationship such as strong and weak magnitude or negative and positive direction. In fact, GWLR has the proficiency of presenting spatially varying regression coefficient values and strength of relationship for each neighbourhood in BMR on a map. Neighbourhoods were chosen as the geographic element of analysis and mapping. Specifically, the study is designed to evaluate the performance and predictive capability of GWLR model of travel behaviour and built environment interactions so as to compare the fitness and outcomes with those of Global Logistic Regression (GLR) model.

The rest of the article is categorized into Section 2 which focused on the background literature; Section 3 description of the data and study area; Section 4 explains the methodology of the study; Section 5 presented and discussed the empirical modelling results, Section 6 focused on the discussion of the implication of the results and in section 7 the study was concluded.

2. Background literature

Several studies have been dedicated to the investigation of the interaction between travel behaviour and built environment characteristics. The most prominent technique for modelling this relationship is the GLR models (Antipova, Wang, & Wilmot, 2011; Haybatollahi, Czepkiewicz, Laatikainen, & Kytta, 2015; Kim & Wang, 2015; Sun et al., 2017; Taaffe, Gauthier, & O'Kelly, 1996; Vega & Reynolds-Feighan, 2009; Vovsha, 1997). This modelling technique is designed to return fixed parameter estimates which are used exclusively to explain the relationship between the built environment, individual characteristics and travel behaviour. Even though the GLR models utilize disaggregate data (Ben-Akiva & Bierlaire, 1999), it results are often aggregated using one value to explain the relationship between variables of an entire region with varying environmental and social characteristics. For example, people often choose different neighbourhood environment to reside depending on their personal and household characteristics with particular concern to the distance they are eager to commute for work purposes. Thus, groups of people who live in different places with diverse attributes associated with their social, demographic and physical characteristics, can exhibit different travel behaviours (Pitombo, Kawamoto, & Sousa, 2011). Travel behaviour has been argued to be influenced by a diversity of factors and such relationship may strongly vary in space.

Previous literature suggests that these factors are either built environment related (Banister, 1997; Boarnet & Crane, 2001; Cao, 2014; Dieleman, Dijst, & Burghouwt, 2002; Ewing & Cervero, 2010; Zhang, Hong, Nasri, & Shen, 2012) or individually oriented such as demographic and socioeconomic (Bhat & Koppelman, 1993; Cheng, Bi, Chen, & Li, 2013; Pitombo et al., 2011; Xiong & Zhang, 2014). It is certainly equivocal why the estimated impact varies with places and regions and how the variables of differential regional policies on land use would likely change commuters travel behaviour. Such knowledge gap (as noted by Zhang et al., 2012), has made it tough for decision-makers to appraise land use policies and plans according to their impact on travel behaviour, and thus, their effect on congestion mitigation, greenhouse gas emission reduction and energy conservation.

GLR is the most common technique in the built environment and travel behaviour interaction literature. For example, choice models (discrete) have recently become increasingly appealing to transport geographers for the study of travel behaviour applications such as mode choice and travel time choice (Ben-Akiva & Bierlaire, 1999; Bierlaire, 2006; Pike & Lubell, 2016; Rahul & Verma, 2017; Sun et al., 2017; Varotto, Glerum, Stathopoulos, Bierlaire, & Longo, 2017; Vovsha, 1997; Zhang et al., 2017). The major advantages of GLR models are that they have the ability to account for parameter and error estimation since the aggregation within the data group has been taken into consideration. In addition, the model's outstanding performance depends on the identification and categorization of the key influencing factors (Zhang et al., 2017).

Interestingly, GWLR models extend the benefits of global models two steps further by strengthening of fixed parameters across space. GWLR can capture potential heterogeneity in the process of measuring the relationship between travel behaviour and the highlighted explanatory variables. Furthermore, the model can display the parameter estimates and strength of relationship on a map, depicting spatial variation by neighbourhoods. However, one debatable assumption of the GLR models is that the impact of explanatory factors on travel behaviour is stationary across neighbourhoods.

To account for the presence of potential spatial non-stationarity in the process of producing the geographic data several empirical spatial models which have emerged in recent years, and have been appropriately applied in diverse areas of interest. Some examples of these applications are the Brunsdon, Fotheringham and Charlton Geographically Weighted Regression (GWR) method (Du & Mulley, 2012; Fernandez, Chuvieco, & Koutsias, 2013; Mathews & Yang, 2012; Nkeki & Osirike, 2013; Pirdavani, Bellemans, Brijs, Kochan, & Wets, 2014; Rodrigues, de la Riva, & Fotheringham, 2014; Selby & Kockelman, 2013; Wang & Chen, 2017; Zhao, Chow, Li, & Liu, 2005), Bayesian Poisson models (Abdel-Aty, Lee, Siddiqui, & Choi, 2013; Aguero-Valverde, 2013; Aguero-Valverde & Jovanis, 2008; Huang et al., 2016; Lee, Abdel-Aty, & Jiang, 2014), autologistic models (Augustin, Mugglestone, & Buckland, 1996; Flahaut, 2004), and Geographically Weighted Poisson Regression (GWPR) technique (Hadayeghi, Shalaby, & Persaud, 2010; Nakaya, Fotheringham, Brunsdon, & Charlton, 2005).

GWR is the widely adopted spatial modelling technique. This is because it outperforms global models and disaggregates results into geographic units. Few or no studies have used such spatial parameters to explain the spatial variability in travel behaviour and built environment interaction. GWR is a semiparametric Gaussian error term designed for estimating numerical responses (Nakaya, 2014). However, the response data for travel behaviour (such as mode choice) is often discrete (dichotomous) in format and such can appropriately be analysed using a spatially-based logistic model. The GWLR model that is applied in this study is a modified extension of GWR specifically designed to fit binary outcome with geographically varying coefficients (Fotheringham, Brunsdon, & Charlton, 2002; Nakaya, 2014).

In summary, it is clear in previous literature that GWR and other spatially weighted models outperforms global models for numerical and count responses (Du & Mulley, 2012; Hadayeghi et al., 2010; Huang et al., 2016; Nkeki & Osirike, 2013; Wang & Chen, 2017). Though limited studies have analysed discrete responses using appropriate spatial models such as GWLR, the conclusion still shows that such models estimation returned better fitness than global models. For instance, Wang, Kockelman, and Wang (2011) used combined multinomial logit statistics with GWR approach to anticipate five classes of land use change in Austin, Texas, and control for some built environment characteristics. Their result suggested that the multinomial logit-based GWR model worked reasonably well with their discrete response datasets. Furthermore, Paez (2006) used a binomial probit model which was weighted geographically to model spatial variability using data from California's BART system. The results show that considerable parametric variation exists across geographical space and that statistical fitness of the local models, was found, by means of a likelihood ratio test, to be higher than the global (homoscedastic) model.

2.1. Geographically weighted regression

The model GWR was introduced by Brunsdon et al. (1996); Fotheringham, Charlton, and Brunsdon (1998) and Fotheringham et al. (2002). They used the term to describe a family of spatially derived regression models that are designed to assign weight to the observations in a dataset which is dependent on the distance from a particular geographic location referred to as regression point. The conceptualization of GWR is to model spatial data so as to understand spatial processes and this was achieved using the concept of local likelihood (Fotheringham et al., 2002). This technique is essentially designed to explore process heterogeneity in the spatial dataset. Spatial non-stationarity describes a scenario whereby global models cannot properly explain the relationship between variables (Brunsdon et al., 1996). Like some other local modelling techniques, GWR seeks to model spatial process heterogeneity found in geographic datasets. It provides regression estimates at every point within the sample region and also points that may not have been sampled in the same region (Wolf, Oshan, & Fotheringham, 2017).

The weight assigned to each observation in the data is based on distance-decay conception in which the weight of the observations reduces as they move farther away. This distance-decay weighting system operates by a kernel function mechanism designed to reduce the influence of the farther observations on the location of interest. The GWR method uses the Ordinary Least Square (OLS) regression model's equation (see Eq. (1)) in its initial development. The fundamental difference is that unlike the OLS the GWR take into cognizance the geographic characteristics of the dataset by incorporating the location coordinates of each data points into the equation (see Eq. (3)). To properly comprehend the development of the GWR model, considering the global regression formulation is paramount. It is written as:

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \tag{1}$$

where y_i denotes ith observation of the criterion variable as it related to each of the independent variable(s) x_k, the beta signs (β_0 to β_k) represent the number of coefficients of the predictors to be estimated, x_{ik} is the observation of the corresponding kth independent variables in the model, while ε_i represents the error term with zero means. If these conditions are satisfied, the OLS parameter to be estimated using the matrix notation, is written as:

$$\hat{\beta} = (x^t x)^{-1} x^t y \tag{2}$$

where x is a matrix of the independent observations with the elements of the first column set to 1, y denotes the vector of the dependent observations, and $\hat{\beta}$ denotes the vector of the OLS coefficients to be estimated, while $(x^t x)^{-1}$ is the inverse of the variance matrix.

The GWR model extends the OLS equation by permitting local parameters rather than global ones to be estimated (Fotheringham et al., 2002). The model which assumes non-stationarity in the process of exploring relationships produces an equation for all the components in the data by standardizing each one using the target feature and its corresponding neighbours. It is designed to consider the spatial element in a dataset by integrating the geographic coordinates of each observation in the equation. The precept behind the GWR technique is that parameters are estimated locally (that is anywhere within the region of interest) given a dependent variable and independent variables which are often measured at any known location i (Brunsdon et al., 1996). To disaggregate the parameters of the global model, Brunsdon et al. (1996) and Fotheringham et al. (2002) rewrote Eq. (1) (by incorporating geographic coordinates) as:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \tag{3}$$

where (u_i, v_i) is a spatial element depicting the geographic location of point i within the region of interest, $\beta_k(u_i, v_i)$ is the value of the kth parameter at location i. In otherwords, it is the coefficient of the independent variable(s) at point i. This parameter value is often presented and measured in a smooth surface allowing certain points to indicate the spatial variability of the surface (Fotheringham et al., 2002).

Eq. (3) simply shows that for each geographic location (u_i, v_i), the value of the criterion variable y_i is explained by the changes in the parameter estimates or coefficients β_k predictors x_k. Unlike the OLS model that absolutely generalizes the parameter estimate of the predictors at a point i where the data was collected as shown in Eq. (1), the GWR compute the parameter estimates of each independent variables for locations that lie between data points and this ease up the generation of a detailed smoothing map of the predicted spatial variations in relationships (Brunsdon et al., 1996). In GWR, estimating the parameters requires extending the matrix notation in Eq. (2) (which uses a constant weighting method) in such a way that each observation is weighted according to its closeness to location i. The formulation by Fotheringham et al. (2002) and Brunsdon et al. (1996) basically introduced a spatial component and distance-decay concept in the weighting which is written as:

$$\hat{\beta}(u_i, v_i) = (x^t w(u_i, v_i) x)^{-1} x^t w(u_i, v_i) y \tag{4}$$

where $\hat{\beta}$ denotes an estimated β, and $w(u_i, v_i)$ represents a n by n matrix where each observed point i in the region under study is geographically weighted. Hence, Eq. (2) was rewritten as:

$$\hat{\beta}(i) = (x^t w(i) x)^{-1} x^t w(i) y \tag{5}$$

where i is the location within the matrix element, while $w(i)$ is a weighted structure that is based on the closeness of point i to the sampling locations around i (capturing the variation of the weight w with location i).

3. Study area and data preparation

To establish the travel behaviour and built environment interaction, several types of data sets were used in this study such were grouped into spatial and non-spatial data. Most of the spatial datasets were utilised in defining the built environment (which describes the physical results of land use planning and capital investments (Wang & Chen, 2017) of the

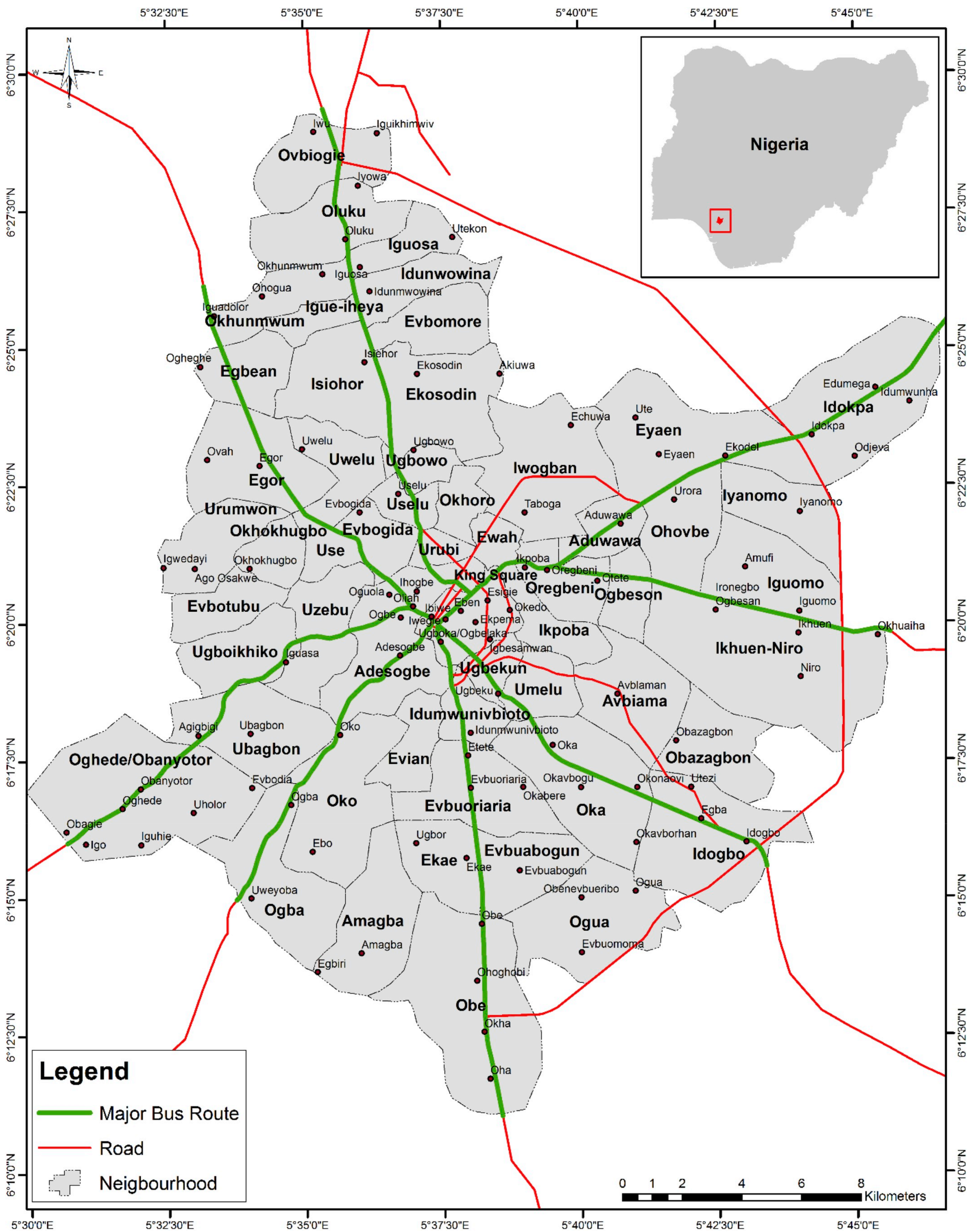


Fig. 1. The 55 delineated neighbourhoods of Benin Metropolitan Region. Source: Updated by Authors from Connah (1975).

study area while the non-spatial datasets described individual attributes of household and travel mode characteristics (e.g. demographic and socioeconomic datasets).

3.1. Study area

The data for this study is based on datasets available for Benin metropolis one of the most vibrant traditional cities in Nigeria and a fast-growing commercial area, supported by government and financial institutions with over 50 identified major residential neighbourhoods. The area is located geographically between latitudes 6° 16′ to 6° 33′ N of Equator and longitudes 5° 31′ to 5° 45′ E of Greenwich Meridian and serve the role of administrative capital of Edo State (see Fig. 1).

Right from the outset, BMR had always been a region of attraction because of its commercial and administrative roles. For example, the ancient Benin Empire was prominent and regarded as the centre for trade in ivory, pepper and slaves. The kingdom's artisans were noted for wood, ivory carving and bronze casting. These socioeconomic roles are still functional (though gradually fading away) in this contemporary era. Presently, BMR has experienced a transformation from agro-based socioeconomic activities to a growth pole of commercial and administrative functionality, supported by numerous financial establishments, educational, health and other plethora of corporate activities. Basically, the urban economy is dominated by government in the formal sector and trading in the informal sector.

The region is characterised by heterogeneous population consisting of the indigenes and migrants from other parts of the country. Originally, it was the hometown of the Binis or Edos. The region has witnessed a tremendous growth in population (Onokerhoraye, 1977) between 1952 and 2006. By 1952, the population of Benin was 53,753 and in 2006, the population for the region rose to 1,085,676 (NPC, 2006). Estimate from census statistics shows that the region is composed of 248,620 households and an average of 6–7 persons per household.

BMR is well-connected by road, though the region has poor quality road structure and related infrastructure which has over the years inhibited accessibility and free traffic flow. The urban core is accessible from every direction of the metropolitan region. This is very much tied to the fact that access roads to such axis are mostly dual carriageways. These roads are fairly well designed and have recently been upgraded by constructing sidewalk on both sections of each road. However, during peak hours, the expected ease in accessing the urban core or the CBD is critically impeded by traffic congestion. This is owing to the stark reality that an enormous amount of vehicles must access such location within the time frame (peak hours).

The metropolitan region generally depends on vehicular mode for urban travel, specifically, work-related travel. Unlike cities of most developed countries, the region lacks effective means of mass transit such as intracity rail or the modern speed rail system, intracity inland water transport, effective bus services, etc. The major modes of urban travel in the region are: private car, taxi, public minibuses (locally called *tuketuke*), medium compartment bus (known as *comrade bus*), small compartment bus (such as *uniben shuttle bus*) and recently introduced engine-based tricycle (popularly called *keke*). The uniben shuttle buses were originally designed to provide cheap transportation services for the students of the University of Benin. These buses ply two major routes within the metropolitan region-Ugbowo-Ring Road route and Ugbowo-New Benin Market routes. Recent government policy action took the tricycle off the major roads and restricted it to residential access roads. The comrade bus is a joint government and private partnership urban transit initiative that ply's the trunk roads connecting the CBD (Ring Road axis) with the suburban areas of the region. Basically, the taxis and tuketuke buses are not tailored to any custom route as they are ubiquitous across almost every route within the metropolitan region.

The public transport is uncoordinated and mainly in the hand of individuals. Most often personal cars are converted to public transport so as to increase household earnings. The region like most cities in Nigeria is characterised by very low household earnings and income. For example, a recently conducted survey (Nkeki, 2018) revealed that over 61 percent of the households in the region earn less than ₦50,000 monthly (i.e. less than $130).

3.2. Spatial data

The built environment data was mined from the spatial data such as Shuttle Radar Topography Mission (SRTM) elevation data of the region with a resolution of 30 m and this was downloaded from NASA's website through earth explorer platform. In addition, a high resolution (2 m spatial resolution) Worldview satellite imagery of the City of Benin in multispectral modes (captured on the 22nd December 2015 with 0.00% cloud cover) were acquired from Digital Globe foundation, Colorado. Road network (vector dataset) of the City was mined from open street map 2016 database. These datasets were manipulated and analysed using various GIS and spatial statistical techniques to generate data for indexes adopted to define the built environment variables.

For survey and spatial statistical analysis purposes, neighbourhood-level data was generated within the GIS system. This produced a consolidated shapefile of the neighbourhood boundary of the region. 55 of these neighbourhoods were identified in the generality of the region (see Fig. 1). This was delineated based on the prominent traditional quarters of ancient Benin City as presented by Connah (1975). He delineated the boundaries of Benin City quarters using the ancient linear earthworks (moat and walls). Other areas that were not covered by Connah's delineation (particularly, the northeast part of the metropolis and most parts of the city of the edge) were updated from the field with the aid of map archive from the Ministry of Physical Planning and Urban Development and the Federal Surveys of Nigeria. Neighbourhood was adopted as the base units of analysis for two reasons: it presents a fine grain and micro-scale geographic units which will provide better homogeneity in analysis and result; the city has no properly delineated transportation planning structure or traffic analysis zone system, existing geographic units are politically delineated e.g. Local Government Area (LGA), geopolitical zone etc.

3.3. Non-spatial data

The non-spatial datasets were collected from primary source following a personal household travel survey conducted in the last quarter of 2017. The datasets which define the individual travel characteristics of the households of the region was collected with the aid of questionnaire and the questionnaire was divided into five parts: The first part includes the household socioeconomic characteristics; the second part is comprising questions relating to the participant's demographic characteristics; the third part includes questions on the respondent's lifestyle; the fourth includes questions on the residential location and other neighbourhood characteristics; the fifth was based on travel-related activities.

Using the central limit theorem sample size estimation as prescribed by Lenth (2011) with 0.03 (3%) margin of error and 99% confidence level, a total of 1,830 households were selected from the region's estimated total of 248,621 (NPC, 2006). Residential houses were systematically selected from each of the 55 delineated neighbourhoods and their respective GPS coordinates were captured using smartphone devices. Neighbourhoods with high, medium and low urban density were assigned 36, 32 and 30 questionnaires respectively. The home-based survey which took roughly 3 months to complete, was conducted by the researcher and 5 well-trained research assistants with a minimum qualification of a national diploma. Out of a total of 1,836 questionnaires administered, 1,736 were considered valid.

Another source of non-spatial data is a checklist specially designed to obtain land use information (with specific reference to residential and commercial types) at the neighbourhood level. This involves the counting of residential, employment land uses and a mix of both. This was carried out with the aid of 200 m × 200 m quadrat covering about 40,000 m^2 of land. Ultimately, the delineated 55 neighbourhoods were used for this data collection. The centre point of each neighbourhood (polygon) was determined within the ArcGIS environment using the mean centre workflow found in the spatial statistics tool of the Arc toolbox. This tool statistically determined the mid-point of each polygon and superimposed point feature on the locations. Using this centre feature, a 200 m × 200 m quadrat was buffered out and with the aid of the worldview satellite image, the locations on the ground were identified prior to fieldwork. Progressively, the number of residential dwellings, employment location and a mix of both were recorded for the 55 locations from the field. Such data was employed for land use mix (diversity) analysis.

3.4. Dependent variable: travel mode choice

In this study, travel mode choices were aggregated into private and public modal characteristics. Data used here is of non-spatial type and this was extracted from the questionnaire data. However, five mode types were initially identified in the region and were entered into the survey. These include a walk on foot, tricycle, bus, taxi and car.

Fig. 2 presents the percentage share of the number of commuters that choose a particular mode for work travel at the neighbourhood–level. Generally, walking as a mode of getting to work seems to be preferred by many commuters since it returned the highest percentage of usage in most of the neighbourhoods. Region–wise, walking returned a 40.6 per cent usage value, this is followed by bus which returned 27.1 per cent usage value. Car returned 23.4 per cent usage value, taxi returned 4.5 per cent and tricycle returned 4.4 per cent. This means that about 40 per cent of the commuters travel to work on foot, about 27 per cent use the public bus, about 23 per cent depends on the private car, while taxi and tricycle have a corresponding patronage level of 4.5 and 4.4 per cent.

3.5. Explanatory variables: built environment, demographic and socioeconomic characteristics

To generate data for some explanatory variables which helped to define the built environment of the region, indexes of urban density, urban sprawl, urban design and land use diversity were first computed. Computing urban density and sprawl indexes, urban patches were extracted from worldview satellite image by means of GIS and remote sensing methodologies. The worldview image was subjected to various photogrammetric techniques including land use classification.

Essentially, two types of image classification exist in GIS operations-unsupervised and supervised classification. Both procedures were implemented in the image classification because it would yield a highly accurate result. The unsupervised classification with ISODATA clustering approach was applied, this allowed preliminary pixel class assortment and interpretation.

Performing a supervised classification, training samples were extracted with the aid of the region of interest tool (ROI) of ENVI software and these sample polygons were stored as a spectral signature file to be used in the classification analysis. The spectral data collected from the unsupervised classification facilitated the supervised classification procedure. The land use supervised classification was executed adopting the maximum likelihood probability technique. The classified land use data was exported and entered into ArcGIS 10.4.1 software from where such data was manipulated using the raster calculator in the ArcGIS toolbox. The raster calculator was used to extract the built-up area from the classified raster. This facilitated statistical estimation and disaggregation of the patches. The extracted urban patches were used as data for calculating the urban density and sprawl indexes.

Urban growth is characterised by a complex multiplicity of changing geographic dimensions (Taubenbock, Wegmann, Roth, Mehi, & Dech, 2009). Growth either takes a radial pattern to build a huge concentric agglomeration (monocentric configuration) or progress into manifold centroid (polycentric configuration). Notwithstanding, urban growth tends to advance towards the suburban area either by increasing density or by sprawling. Based on evidence from previous literature (Comendador, López-Lambasb, & Monzónb, 2014; Crane & Chatman, 2003), density or compaction may reduce travel by car and encourage non-motorised mode. Sprawl may initiate long-distance travel, encouraging motorised travel particularly public mode, if such sprawl occurs along trunk roads or may encourage private car mode if sprawl is haphazard.

Urban density as an indicator of the built environment was quantified using the patch density index (PD) spatial metrics. There are two major techniques of computing urban density, these are patch density and housing density. In this study patch density was preferred because urban density is better measured using urban patches which consist of buildings, roads and all built up surfaces and landscape within the urban area. Housing density would only capture the residential density of the area leaving other as much important constituent of the urban landscape out of the data. Values of PD are the number of urban patches of the conforming patch type divided by total landscape area in m^2. This was computed using the formula presented by McGarigal and Ene (2014):

$$PD = \frac{N}{A}(1000,000) \tag{6}$$

where:

N = aggregate number of patches in the landscape, excluding any background patches;
A = total area of landscape in m^2 (the value of PD was multiply by 1000,000 to convert to km^2 so as to interpret the PD based on number of patches per km^2). However, given any value of PD indicates the number of patch per km^2. PD increases as the urban landscape continues to disaggregate and less compacted, while it decreases as the urban landscape becomes compacted. FRAGSTATS 4.2.5 was used for this computation because of its prominence and widely used for urban landscape characterisation (Deng, Wang, Hong, & Qi, 2009; Nkeki, 2016; Ramachandra, Aithal, & Sanna, 2012; Taubenbock et al., 2009).

Urban sprawl is another important factor for measuring built environment. Quantifying urban sprawl has been a long time issue in academic research. The formulation of empirical indicators has made the burden of measuring sprawl lighter. Today, entropy index in its various modification has become a dependable means of characterising urban sprawl (Bhatta, Saraswati, & Bandyopadhyay, 2010; Nkeki, 2016; Sarvestani, Ibrahim, & Kanaroglou, 2011; Sudhira, Ramachandra, Raj, & Jagadish, 2004), this was made possible by the recent improvement in remote sensing and GIS. In this study, Shannon's entropy index was calculated to quantify urban sprawl manifestation by determining the magnitude of concentration of patches within the various neighbourhoods. Entropy statistic was computed for the study region based on the neighbourhood-level. The result of the entropy statistic was then used as values for urban sprawl variable under the built environment factor.

Shannon's entropy index for urban sprawl was computed here with the formulation by Bhatta et al. (2010):

$$H_n = -\sum P_i In(P_i) \tag{7}$$

where:

H_n = Shannon's entropy index;
P_i = proportion of built-up patches i in each neighbourhood;
n = aggregate number of neighbourhoods in the region. Shannon's entropy index varies from 0 to $In(n)$, and indicates compactness of

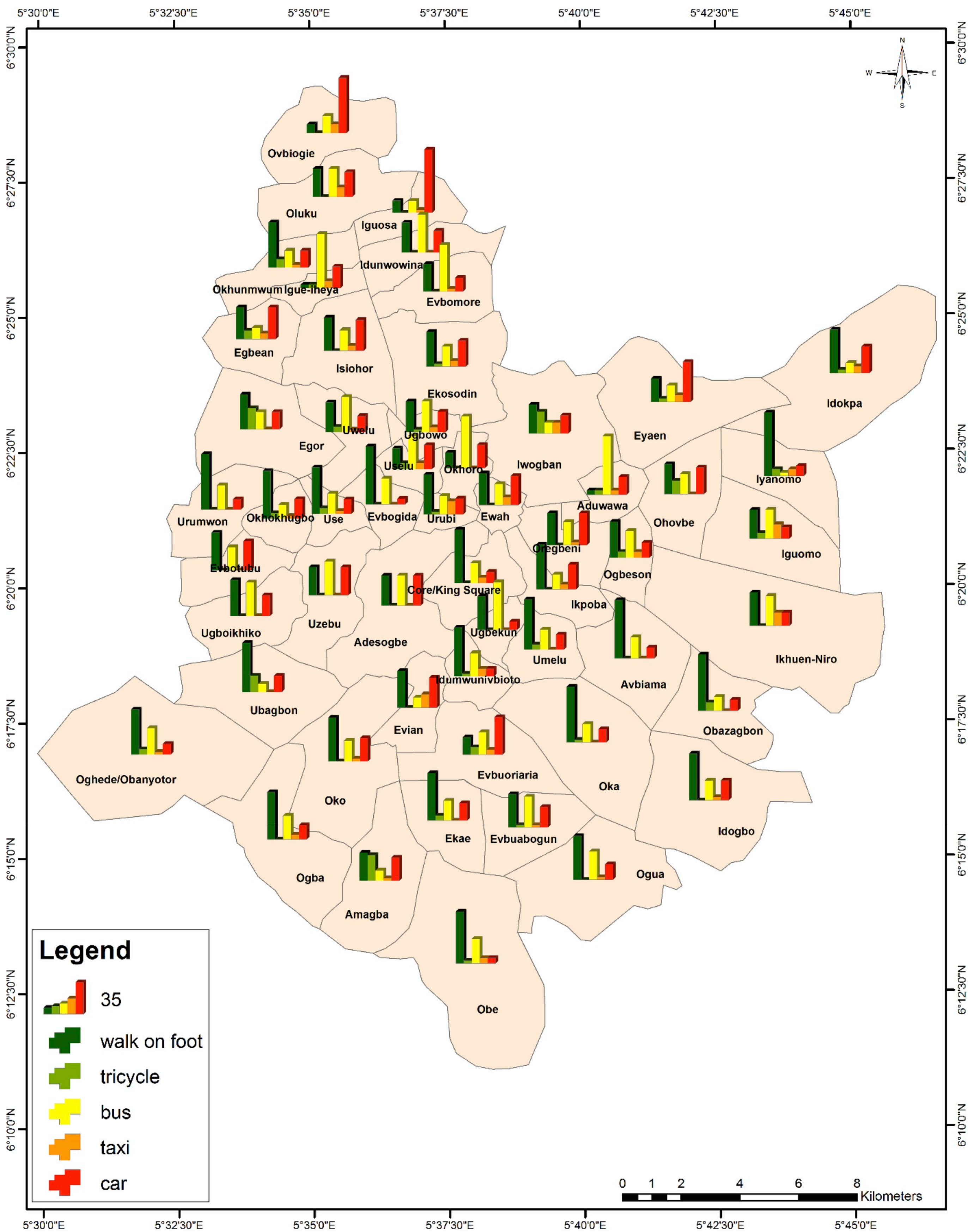


Fig. 2. Percentage of travel mode choice by neighbourhoods.

urban patches for values near to 0 and dispersed distribution for values near to $In(n)$. Entropy values larger or near to $In(n)$ depicts dispersion of urban patches or urban lands which is interpreted as the occurrence of urban sprawl. This spatial index was calculated using Eq. (7) which was entered into the field calculator of ArcGIS and python platform was used to perform the computation. The data for this analysis is the urban patch features that were mined from the land cover classification.

Land use mix (entropy) was calculated using data from the 200 m × 200 m quadrat consisting of the number of residential dwellings, employment location and a mix of both which was recorded for the 55 neighbourhoods in the region. The purpose of modelling the main land use types is to generate an additional variable which was used for defining the urban built environment factor. The land use mix analysis was modelled after Shannon's entropy index. However, the entropy index has become the generally used technique of quantifying land use mix and diversity in contemporary literature. For example, it has been used to determine the level of homogeneity or diversity of various land uses, such as employment, residential land uses (Strauss & Miranda-Moreno, 2013; Zhang et al., 2012). The entropy index after Zahabi, Miranda-Moreno, Patterson, and Barla (2012) is defined as:

$$E_j = (-1) \times \sum_j \frac{P_j\, In\,(P_j)}{In\,(J)} \tag{8}$$

where:

E_j = land use mix entropy index
P_j = proportion of land use i in neighbourhood j
j = number of varying land use type in the neighbourhood

In this study, j = 2: residential and employment land uses. The value of E_j varies between 0 and 1. 0 corresponds with homogenous land use characterised by a single land use while 1 refers to a perfect mix in which all land use types are represented equally. Eq. (8) was briefly modified (Bahadure & Kotharkar, 2015) to allow computation of a 2 category mix or diversity types (commercial and residential). The equation is defined as:

$$Land\ use\ mix\ (EI) = (-1) \times \frac{\left[\left(\frac{b1}{a}\right) \times In\left(\frac{b1}{a}\right) + \left(\frac{b2}{a}\right) \times In\left(\frac{b2}{a}\right)\right]}{In\,(n)} \tag{9}$$

where:

EI = land use mix entropy index
a = aggregate number of land uses of the two land use types
$b1$ = the commercial land use type
$b2$ = the residential land use type
n = the aggregate number of land uses in the mix (in this case 2)

However, the calculation of this index was done with the field calculator algorithm in ArcGIS software at the neighbourhood-level using Eq. (9).

Urban design as an indicator of the built environment and explanatory variable for this study has been characterised specifically by site design, block size, dwelling and street characteristics. Prominent among these is the use of street network features within an area (Ewing & Cervero, 2010). In this study, urban design was quantified using a four-way street intersection points extracted from the worldview satellite imagery. This was used to evaluate the number of street intersections per square kilometre in each neighbourhood. To analyse the geographic point data, a Street Intersection Density (SI Density) index was developed and it is presented as:

$$SIDensity = \frac{NoN}{A}\,(1000{,}000) \tag{10}$$

where:

$SIDensity$ = street intersection density
NoN = number of nodes (point of street intersections in a particular neighbourhood)
A= total area of the neighbourhood in m^2 (the value of SI Density was multiply by 1000,000 to convert to km^2 so as to interpret the index based on number of nodes per km^2). It is believed that where intersections are more prevalent, travel tend to be shorter (Reiff, 2003). The density of the four-way intersection helps to define the level of grid design that a particular neighbourhood is structured. A grid street design may encourage short distance travel; it may also discourage public mode of travel. This is because transit performance is effective on a long distance and mostly corridor road. The SI Density index was computed in ArcGIS from where the intersection points were extracted as vector format.

Other explanatory variables used to define built environment for this study are distance of neighbourhood to CBD; transit accessibility; availability of sidewalk in the neighbourhood and the number of parking spaces within the neighbourhood. Distance to CBD was defined using the average distance of the neighbourhood to the CBD through a regular bus route or a major road in such neighbourhood. Transit accessibility was measured using field data on the number of regular bus route within the neighbourhood. The data for the last two variables were mined from the questionnaire survey. Data for demographic and socioeconomic variables were mined from the questionnaire forms and a description of these variables are highlighted in Table 1.

4. Methodology

A multivariate factor analysis technique was first computed as a data reduction procedure of the dataset and retained only the significant factors of built environment attribute, socioeconomic and demographic characteristics. This was ascertained with both the positive and negative high factor loadings. The principal component extraction method was used to extract the sum of squared loadings and the oblimin factor rotation method with Kaiser normalisation was adopted for the analysis. The components with eigenvalues above 1 were retained as the extracted significant variables. Since categorical variables are not appropriate for factor analysis, all such variables in Table 1 were standardized before carrying out the analysis. The method used for the transformation of the categorical variables is the nonlinear optimal transformation (Meulman, 1992) which has the capability of assigning quantitative values to qualitative scales. The statistical package for social sciences (SPSS) was used to conduct the analysis. Within the SPSS software an extension or plug-in known as CATPCA (categorical principal components analysis) was used for the binary, ordinal and nominal variable types transformation. The significant factors were then extracted as independent variables to be used for GWLR model.

GWLR was used to explore the correlates between the explanatory variables consisting of built environment characteristics, aggregated household attributes and the dependent variable which is the travel mode choice at the neighbourhood-level. The travel mode choices were aggregated into private and public modal characteristics. The GWLR is a modified extension of geographically weighted regression (GWR). The latter is a Gaussian error term suitable for modelling numerical responses for all variables. With respect to modelling count or binary (categorical or dichotomous) responses, like other global model types of generalised linear modelling, particularly logistic and Poisson regression, the GWLR is a semi-parametric and a natural extension of GWR designed to theoretically derive geographically weighted generalised linear models.

GWLR can be used for modelling binary dependent variable, respectively, with geographically varying coefficients using both local and global terms (Nakaya, 2014). GWLR is a local spatial statistical

Table 1
Definition of explanatory variables for GWLR.

Variables	Description
Demographic:	
1 Gender	0 = male; 1 = female (nominal variable)
2 Age of respondent	1 = below 20; 2 = 21–40; 3 = 41–60; 4 = above 61 (ordinal variable)
3 Household size	Continuous variable
4 Marital status	1 = single; 2 = married; 3 = divorce (nominal variable)
5 Number of children	Continuous variable
6 Period lived in the neighbourhood	1 = < 1 year; 2 = 1–5 years; 3 = 6–10 years; 4 = 11–15 years; 5 = 16–20 years; 6 = 21 years and above (ordinal variable)
7 Origin	0 = migrant; 1 = indigene (nominal variable)
8 Family orientation	0 = traditional; 1 = modern (nominal variable)
Socioeconomic	
9 Household monthly income[a]	1 = below N 50, 000; 2 = 50,000–69,000; 3 = 70,000–99,000; 4 = 100,000–169,000; 5 = 170,000–199,000; 6 = 200,000–269,000; 7 = 270,000–299,000; 8 = 300,000–369,000; 9 = 370,000–399,000; 10 = 400,000–469,000; 11 = 470,000–499,000; 12 = 500,000 and above per month (ordinal variable).
10 Job type/industry	1 = Farming; 2 = manufacturing; 3 = health; 4 = government/civil service; 5 = transportation; 6 = telecommunication; 7 = finance; 8 = wholesale and retail; 9 = education; 10 = services; 11 = legal and law enforcement; 12 = applicant (nominal variable)
11 Number of jobs in the household	Continuous variable
12 Number of cars per household	Continuous variable
13 Education	1 = no formal education; 2 = primary education; 3 = secondary education; 4 = tertiary education (ordinal variable)
14 Number of driving license in the household	Continuous variable
15 Residential tenure	1 = owner; 2 = rented (nominal variable)
Built environment	
16 Land use mix (diversity)	Value of entropy index (proportion of workplaces to residential places within the neighbourhood). Continuous variable ranging from 0 to 1
17 Urban density	Value of urban patch index (PD) (density of built-up patches in a neighbourhood). Continuous variable
18 Distance to CBD (km)	The average distance of neighbourhood to CBD (continuous variable)
19 Urban design (neighbourhood street intersection density)	Value of street intersection density index (SI Density) by neighbourhood. Continuous Variable
20 Urban sprawl	Value of entropy index. Continuous variable ranging from 0 to 1
21 Availability of sidewalk in the neighbourhood (urban design)	0 = yes; 1 = no
22 Number of parking space within the neighbourhood (urban design)	Continuous variable
23 Transit accessibility	Total number of regular bus route within the neighbourhood (continuous variable)

[a] When the survey was conducted ₦1.00 was roughly equal to $0.0026.

model designed to capture both spatial association and diversity (heterogeneity) simultaneously. This model is frequently referred to as disaggregate statistics (Fotheringham et al., 2002), because it has the ability of spatially disaggregating global statistics into a defined area unit, such as, neighbourhoods which are depicted in the geographic information system (GIS) with polygon or point. Unlike the global statistics, the final computation of the GWLR model yield multi-valued results, such as the parameter estimate, R^2, *p-value*, etc., and this can be mapped in GIS using spatial features (in vector or raster grid format). Although this model disaggregates the outcome of geographic data spatially by presenting point by point or unit by unit (polygon) explanation of the relationships that exist between variables, it uses aggregated data of individual responses especially when samples are drawn from defined area units. For example, the population distribution of a country does not show the actual spatial distribution or location of people but an aggregated spatial distribution of people into defined area unit or boundary.

The GWR in its modification include generalised weighted linear models to allow for use in categorical response analysis. It has gained widespread popularity and has been applied in diverse field of studies. For example, in the health and epidemiological field (Nakaya et al., 2005; Nkeki & Osirike, 2013), in demography (Mathews & Yang, 2012), in accessibility study (Du & Mulley, 2012) in transit ridership modelling (Zhao et al., 2005), ecological disaster (Fernandez et al., 2013; Rodrigues et al., 2014), etc. The development of GWR is discussed in further detail in section 2.

In this study, GWLR was used to show the relationship that exists between the built environment, individual attributes and travel mode choice based on point by point comparison. Fundamentally, a GWLR model is shown as $y_i \sim$ Bernoulli $[p_i]$ meaning approximate binomial distribution which is explained by this equation:

$$logit\,(p_i) = \sum_k \beta_k(u_i,\ v_i)x_{k,i} \tag{11}$$

The dependent variable must be 0 or 1. However, p_i is the modelled probability that the dependent variable becomes 1. Its semi parametric variant is described as:

$$y_i \sim \text{Bernoulli}\ [p_i]$$

$$logit\ (p_i) = \sum_k \beta_k(u_i,\ v_i)x_{k,i} + \sum_I \gamma_I Z_{I,i}$$

Where:

$u_i,\ v_i$ = the local coordinates in space of point *i*
$Z_{I,i}$ = the *I*th independent variable with a fixed coefficient γ_I
$x_{k,i}$ = the explanatory variables.

The last session of Eq. (12) represents the global statistic. Specifically, the GWR/GWLR modeller version 4.0 software was used for the analysis. The dependent variable consists of the travel mode choice of commuters aggregated into the two major modes-private and public. However, the essence of this spatial modelling is in respect to the assumption of this study that analysing individual-level data alone without considering the neighbourhood effect, in which such individual behaviour originate from, may lead to failure to account for the variability between places.

The GWLR was computed using *adaptive kernel* type (*adaptive bi-square*) because it has the ability to change local extent by controlling the *k*th nearest neighbour distance for each regression location. The golden section search option in the GWR software was selected because it can automatically determine the optimized bandwidth size for the data. However, the neighbourhood vector map of BMR which formed part of the data used for analysis was entered into ArcGIS software from where spatial and non-spatial attributes were manipulated. Such manipulation includes entering of the individual attributes, built environment characteristics and travel mode characteristics point data into the corresponding neighbourhood polygon. The x and y coordinates for each point data were automatically computed and loaded into the attribute table within the ArcGIS-ArcMap workspace. In other words, 1,028 data points with their corresponding coordinates were created for all the variables and these were used to compute the GWLR. Such detailed ArcGIS attribute table was then converted to CSV format and entered into the GWR/GWLR modelling software for further computation and analysis. Before then, factor analysis was used as data transformation techniques for the set of predictors to reduce the number of variables into significant factor scores which were then used as explanatory variables for GWLR).

The reason for such process and interoperability of data between ArcGIS and GWR modeller is that ArcGIS spatial statistical toolbox presents GWR extension which can only compute linear GWR with the assumption of linearity of the dependent variable which values must be greater than zero. Put simply, ArcGIS-based GWR tool lacks the ability to compute binary outcomes. While the GWR/GWLR modeller and software presents three major algorithms–GWR-Gaussian model type, GWLR-logistic with binary model type and GWPR (Geographically Weighted Poisson Regression)-Poisson with count model type. The GWLR algorithm is designed to run weighted logit models of binary composition. The weakness of the GWR/GWLR software is that it lacks the platform for data entering, editing and visualisation. Therefore, this study relied on both software for this analysis. Progressively, the results of the GWLR was reentered into the ArcGIS system for visualisation and spatial interpolation of the local parameter estimate, local *t*-value and local R^2. The 1,028 points (result) were interpolated based on their corresponding neighbourhood polygons for neighbourhood-level interpretation.

5. Results

The GWLR model is adopted in this study to explore the intrinsic relationship that may exist between the predictors and travel mode choice, focusing more on the significant spatial varying relationships at the neighbourhood-level which GLR models could not achieve. It is assumed that the neighbourhoods may present distinct pattern of relationship when results are disaggregated. In fact, GWLR has the proficiency of presenting spatially varying regression coefficient values and strength of relationship for each neighbourhood in the region.

5.1. Defining the explanatory variables for GWLR

To conduct the GWLR model the predictors in Table 1 were entered into factor analysis for the purpose of data transformation and data reduction so as to reduce the chance of multicollinearity among predictors. Factor analysis was computed and instructed to retain eigenvalues above 1. Six factors were retained and these explained roughly 62 per cent of the variation in the data (see Table 2).

The result of the analysis was saved as a regression factor scores and was employed for further analysis. For the purpose of factor naming, the oblimin rotation method was used to generate a pattern matrix to guide such a process (see Table 3). Factor analysis pattern matrix shows that factor 1 loads high on the 6 neighbourhood characteristics, factor 2 loads high on 6 demographic characteristics, factor 3 loads high on 4 socioeconomic variables, factor 4 loads high on the 2 accessibility variables, factor 5 loads high on 1 demographic variable (gender) and finally, factor 6 load high on 2 demographic variables (origin and period lived in the neighbourhood). Factor 5 and 6 seem to load high on demographic variables after factor 2 has captured 6 other demographic variables.

However, factor 1 was named neighbourhood characteristics because it loads high on such variables irrespective of the sign. Factor 2 was named demographic characteristics since it loads high on such variables. Factor 3 was named socioeconomic characteristics in the same way, factor 4 was named accessibility characteristics. Others are factor 5 which was named gender because it returned high loading for gender alone and factor 6 was named origin status and duration of stay in the neighbourhood. Factor 5 and 6 are part of the demographic variables but since they were retained as separate component factors, they were named as such. Neighbourhood and accessibility characteristics are interpreted generally as built environment factors even though factor analysis split it into two dimensions.

To show spatial variations in these 6 factors, they were subjected to GIS operation where each of the factors extracted as explanatory variables for GWLR was interpolated. The results of the factor score mapping are presented in Fig. 3. This gives a clear visual pattern and definition of the 6 explanatory variables. For example, factor 1 (neighbourhood attributes) seem to partition BMR into 3 distinct sub-regions: core axis (high negative factor scores) depicted with blue shade in the centre signifying cluster of points in the neighbourhoods with high urban density, low sprawl capability, high street intersection density and high homogenous commercial land use; peripheral area (high positive factor scores) depicted with red shade at the edges of the city representing cluster of points in neighbourhoods with low urban density, higher sprawl capability, low street intersection density and homogenous land use (i.e. high residential); transition zone (low factor scores) depicted with gradient of lighter blue to ligher red shade which formed a cluster between the core axis and the peripheral area having a combination of both sub-region characteristics including high land use mix.

5.2. Validating the Model's appropriateness

The validation and fitness test of the GWLR were performed in two ways firstly by conducting a global logistic regression (GLR) so as to compare the result with that of the GWLR. The major parameters of interest are the R^2, the degree of freedom (DOF) and the deviance, secondly by conducting geographical variability tests of local coefficients to ascertain whether the GWLR is appropriate for the data. To test the latter's, majority of different of criterion for the predictor variables must return a negative value. When the criterion for all predictors returns positive value it simply suggests no spatial variability in terms of the model selection criteria.

Comparing both models with deviance values (Table 4), show that the value is reduced from approximately 485.995 (for GLR model) to 476.020 (for GWLR model). The difference is about 9.975 implying that local models fitness is higher when explaining spatial dataset. GWLR model improved the explaining power of GLR model with about 7.72 per cent (Table 4). This is a high percentage explained value not accounted for by the GLR model. Table 5 shows that of the 6 factors only 2 have a positive difference of criterion. This suggests that there is spatial variability among the neighbourhoods and GWLR is the appropriate model for the dataset.

5.3. GWLR model estimation results

The GWLR model utilised two generalised mode choice types. They include private and public, walking as a mode was removed from this model since it cannot be directly classified under any of this generalised modes. This will have no effect on the model because walking is a unique mode of its own and not of interest here. As the focus of this study is investigating the factors that may explain the mode choice at the public and private level, since the major objective of built

Table 2
Eigenvalues for component scores.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.784[a]	20.798	20.798	4.784	20.798	20.798
2	3.497[a]	15.203	36.001	3.497	15.203	36.001
3	2.417[a]	10.510	46.511	2.417	10.510	46.511
4	1.319[a]	5.737	52.248	1.319	5.737	52.248
5	1.159[a]	5.041	57.289	1.159	5.041	57.289
6	1.026[a]	4.463	61.752	1.026	4.463	61.752
7	0.968	4.209	65.961			
8	0.917	3.988	69.949			
9	0.834	3.627	73.575			
10	0.786	3.416	76.991			
11	0.761	3.308	80.299			
12	0.679	2.950	83.249			
13	0.607	2.639	85.888			
14	0.578	2.515	88.403			
15	0.523	2.272	90.675			
16	0.503	2.187	92.861			
17	0.454	1.974	94.835			
18	0.387	1.683	96.518			
19	0.212	0.922	97.441			
20	0.182	0.790	98.231			
21	0.166	0.720	98.950			
22	0.137	0.594	99.544			
23	0.105	0.456	100.000			

[a] Retained factor components with eigenvalues above 1.

Table 3
Pattern matrix from factor analysis.

Variables	Component					
	1	2	3	4	5	6
Demographic						
Gender					0.840	
Age		0.648				
Marital status		0.737				
Household size		0.832				
No. of children		0.882				
Origin						0.798
Household orientation		−0.511				
No. of jobs in the household		0.496				
Period lived in the neighbourhood						0.500
Socioeconomic						
Household monthly income			0.728			
Education (highest level)			0.518			
Job industry						
No. of cars in the household			0.853			
No. of driver's license in the household			0.861			
Accessibility						
Residential distance to CBD (km)				0.822		
No. of bus route in the neighbourhood				0.842		
Neighbourhood characteristics						
No. of approved parking spaces in the neighbourhood	0.871					
Is there Sidewalk in your neighbourhood	−0.405					
Urban density index	0.947					
Sprawl index	0.948					
SI Density (Urban design)	−0.712					
Land use mix index (Urban design)	−0.690					

Note: Component loadings less than 0.4 are not shown as such were classified low loadings.

environment and travel behaviour studies is to understand how to reduce over-dependence on personal car travel by encouraging mass transit at various levels so as to promote green, healthy and sustainable travel behaviour. However, before these can be achieved, there is need to improve understanding on the factors that may promote undesired travel behaviour such as over dependence on car for regular urban travel. Built environment cannot force people to adopt certain behaviour, though it may encourage it, but perhaps people may respond to changes in different manners depending on what is available/means. For example, those who choose to ride on a bus may do so for 2 major reasons: first, because they actually are comfortable/prefer bus or second, because they cannot afford other preference such as private car due to poor socioeconomic status. There is need to understand people with these kind of behaviour by so doing a more sustaining option can be proposed such as carpooling for car lovers, etc. Modelling built environment-travel behaviour relationship should be multidimensional since human behaviour is complex.

The observation or responses related to walking mode were likewise removed from the GWLR analysis, this reduced the number of data points from 1,736 to 1,028. The R^2 values are shown in Fig. 4 as a spatial smoothing of GWLR model showing the neighbourhoods or areas where the model's prediction and strength of relationship are improved. Importantly, that there is a regional difference in the strength of relationship in the study region. Overall, the R^2 value (0.72) shows a strong significant relationship between travel mode choice and the various factors. At the neighbourhood level, the strength of relationship ranges from 0.62 to 0.72 (this also depicted by the isoline in Fig. 4). This indicates that the model explains between 62 and 72 per cent with a spatial variation of 10 per cent. Indicating that fluctuation in the strength of relationship among neighbourhoods is somewhat lower than expected. The strength of relationship or percentage explained is higher in the north-west part of the region in Ovbiogie, Oluku, Iguosa, Igue-Iheya, etc and gradually reduces towards the south-west part of the region in Egbean, Egor, Urumwon, Evbotubu, etc. While the north-east part returned the lowest percentage explained with corresponding values of 0.63 to 0.62. Thus, this pattern suggests local variation in the relationship. However, the best fits are found clustering in the northern part of the region.

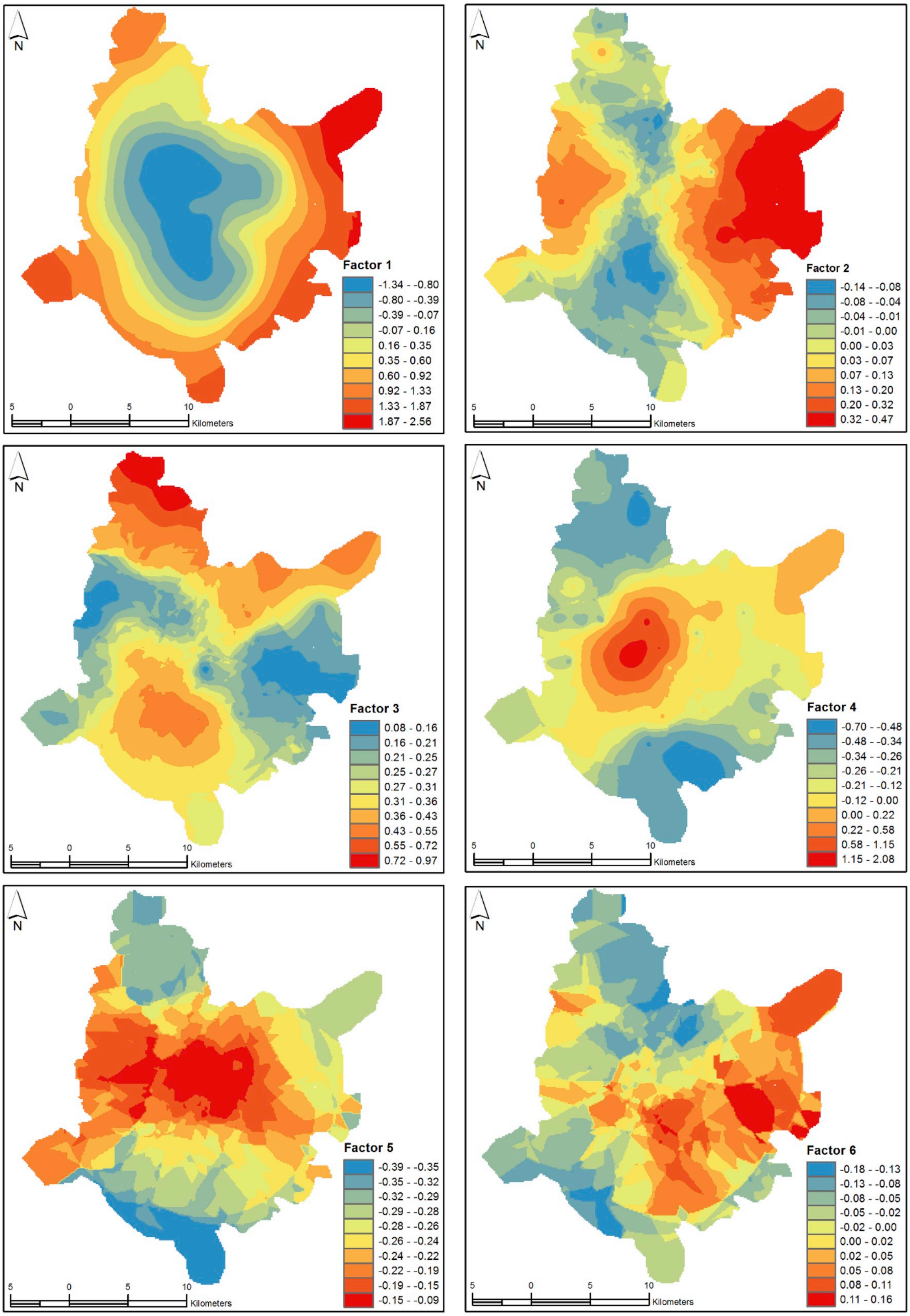


Fig. 3. Explanatory variables for GWLR model.

Accordingly, the higher the local *pseudo-t-statistic* value, the higher the level of significance for that factor regardless of the corresponding sign for such neighbourhood. Significant variables in this spatial empirical model are factors 2 (demographic variables), 3 (socioeconomic variables) and 5 (gender status). The local coefficients and local *t-values* of these three significant factors were interpolated so as to present the results on a continuous raster surface. The interpolation was conducted using 1,028 data points and the neighbourhood polygon data as area

Table 4
Models fitness comparison.

Fitness parameter	GLR	GWLR	Difference
Deviance	485.994640	476.019853	9.974787
R^2	0.647673	0.724904	0.077231
DOF	1021.000	1012.879	8.121

Table 5
Geographical variability tests of local coefficients.

Variable	Diff of deviance	Diff of DOF	Diff of criterion
Intercept	0.352	0.772	1.230
Factor 1	1.118	0.937	0.800
Factor 2	2.396	0.767	−0.824
Factor 3	1.905	0.811	−0.245
Factor 4	1.423	0.224	−0.965
Factor 5	0.401	0.771	1.179
Factor 6	1.623	0.778	−0.030

unit (i.e. the result was aggregated/interpreted at the neighbourhood-level). This generated four pairs of raster surfaces including the intercept of the local constant term for the neighbourhoods.

Fundamentally, the resultant raster surface for the predictors shows that there is spatial variation in the association between travel mode choice and the factors of demographic and socioeconomic attributes. Other factors like built environment consisting of factors 1 and 4; origin status and duration of stay in the neighbourhood were unexpectedly not significant in the spatial model. The local coefficients for the significant factors are displayed in Fig. 5. A positive and negative relationship was shown in the result of GWLR. On the one hand, the positive values of statistical significance imply a direct relationship between the predicting factors and travel mode choice. On the other hand, negative values imply an indirect relationship. Put differently, as the value of the predictor variable increases the chance of choosing a particular mode of travel increases and vice versa.

Fig. 5 presents the local coefficients and their corresponding *t-values* which indicate the significance of the variable at a neighbourhood of interest. The colour ramp for the factor coefficients is graduated from dark to light gold. Neighbourhoods with light shade depict where that particular variable exhibit a strong influence on mode choice while

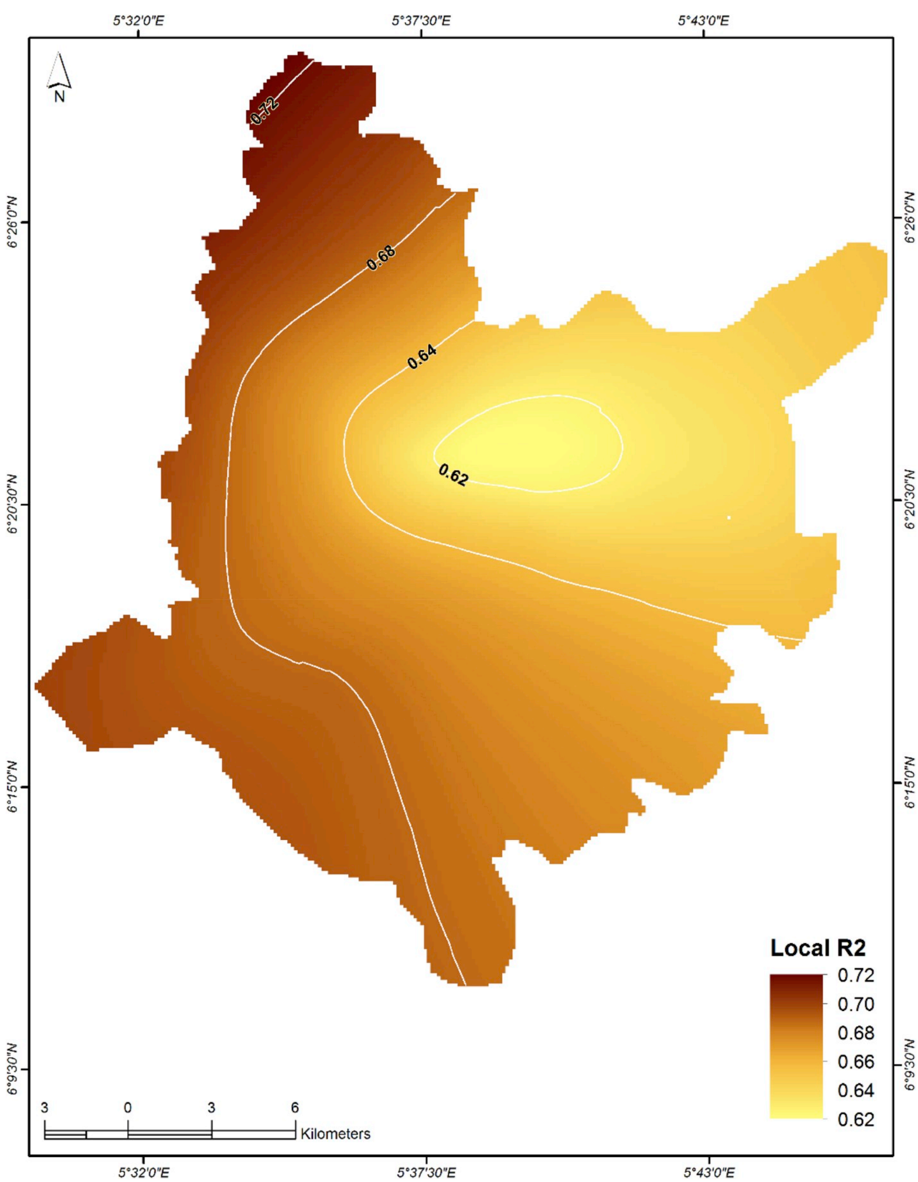
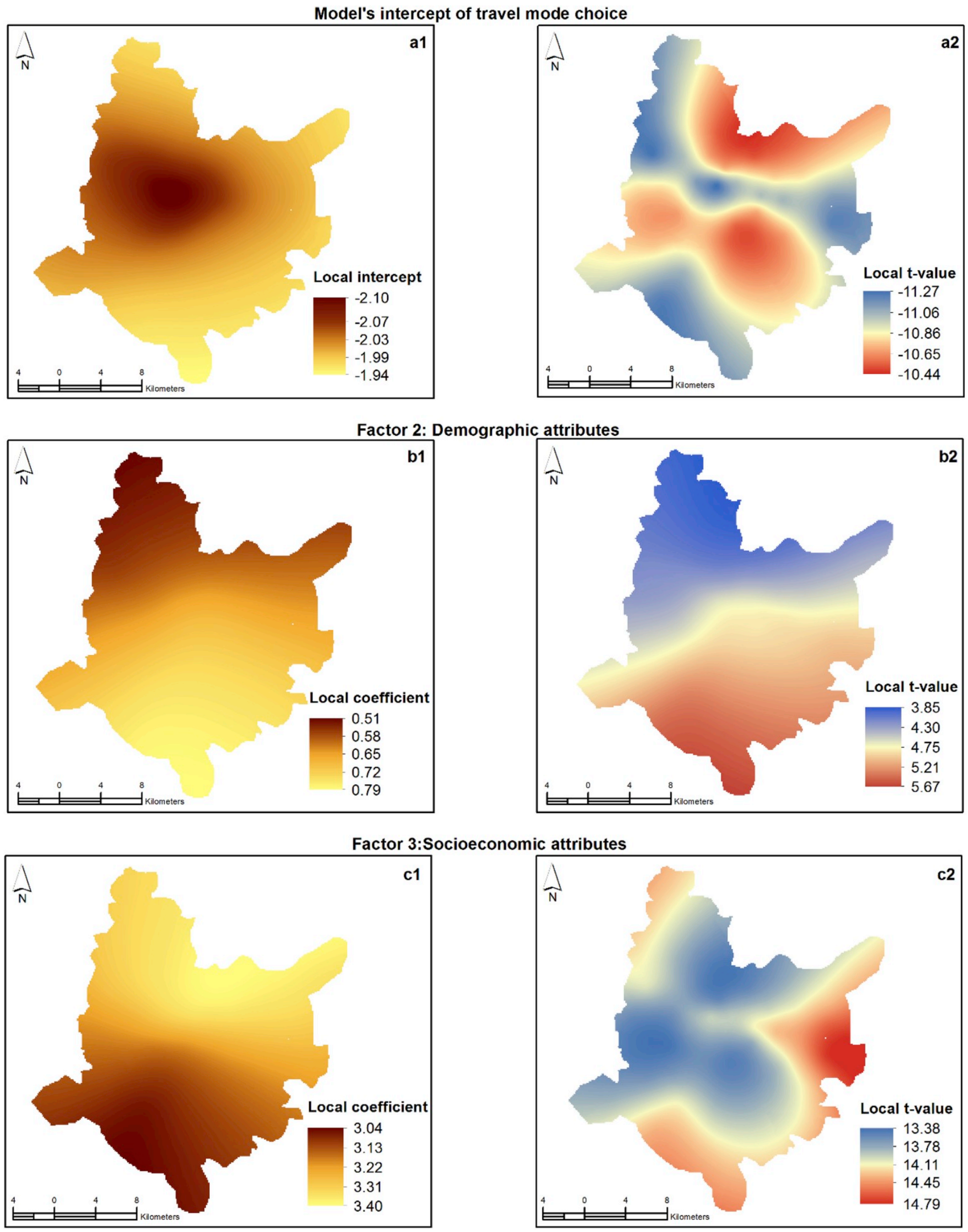


Fig. 4. Local R^2 smoothing for GWLR.



Continue from the next page
Continue from the previous page

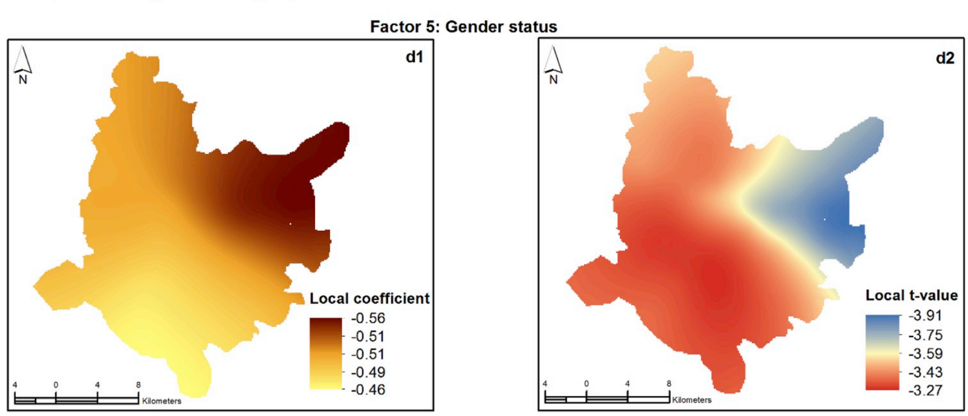


Fig. 5. Local coefficients and local *t*–values (significance) for GWLR model.

dark shade represents neighbourhoods where that specific factor exhibits a weak influence on mode choice.

Demographic variables (Factor 2) seem to exert a strong positive influence on travel mode choices made in the neighbourhoods of the region. This implies that variations in demographic attributes may increase the probability of choosing private mode. The influence is stronger in the southwestern part of the region in such neighbourhoods as Obe, Amagba, Ogba, Ekae, Oko, Evbuabogun, Ogua, Idogbo, Evbuoriaria and this influence gradually increases from the south-western edge of the city towards the city centre and the southeastern periphery. The influence of demographic attributes is very low in the north-west zone of the region and this extends slightly towards the northeastern edge of the city.

Interestingly, Fig. 5 (b1 and c1) revealed a spatial dichotomy between the demographic factor and socioeconomic factor. A critical examination of both local coefficient raster smoothing shows a cluster of neighbourhoods where demographic attributes manifest high influence, the socioeconomic variables manifest low influence vice versa. Unequivocally, GWLR coefficient smoothing shows that socioeconomic attributes of commuters in BMR (Factor 3) exhibit a strong influence on mode choice decision making in the north-east zone of the region (meaning that changes in socioeconomic attributes may increase the probability of selecting private mode of travel). Some of the neighbourhoods under this influence are Isiohor, Egor, Uwelu, Ovbiogie, Evbomore, Idunwowina, Iguosa, Ugbowo, Okhoro, Urubi, Uselu, Igue-Iheya, Ekosodin, Iwogban, Eyaen, Idokpa, etc. Majority of the neighbourhoods in this cluster have higher access to a car. For instance, Iguosa has roughly 70 per cent of it household depending on private car travel (see Fig. 2).

The coefficient for Factor 5 which is gender status returned negative value implying that the association between gender and travel mode choice is inverse i.e. variation in gender may decrease the probability of choosing private mode over public mode. The negative relationship is stronger in the south-west to south-east covering the neighbourhoods where demographic attributes (Factor 2) has a strong positive influence (Fig. 5 b1). This inverse relationship may be the reason why factor analysis extracted gender from demographic variable and retained such as an independent factor score.

6. Discussion

Local mode choice model (GWLR) shows that there is spatial variation in the association between mode choice (private and public modes) and the factor scores of demographic and socioeconomic variables across neighbourhoods. The findings show that built environment variables are not important predictors of mode choice in the region. In the GWLR model, only demographic and socioeconomic variables were found to be significant. By implication, the individual attributes of commuters in BMR are the major factors initiating changes in the travel mode choice decision of workers in the region (Aditjandra, 2013). This finding contradicts the finding of Kim and Wang (2015) which imply that neighbourhood travel behaviour is significantly influenced by neighbourhood characteristics.

It is evident from Fig. 5 that the demographic characteristics (Factor 2) of the households in the region, though heterogeneous, largely drive the choices of travel, i.e. determines whether such household would prefer a private travel mode over public mode. By implication the demographic factor's positive coefficient shows that changes in age, marital status, household size, household orientation and number of jobs may lead to change in mode choice. Simply put, older large size household practicing modern lifestyle with more number of jobs tend to prefer private car mode. The influence which was allowed to vary over space by the local model is higher in the neighbourhoods located towards the south-east and south-west part of the region. Some of these neighbourhoods are Obe, Ogua, Amagba, Ekae, Evbuabogun, Ogba, Oko, Evbuoriaria, Oka, Idogbo, Oghede/Obanyotor, Ubagbon, Ikhuen-Niro, etc. These are neighbourhoods where household size is higher than the regional average. Most likely larger households in the south-east and south-west part of the region tends to depend on private travel mode and families with modern lifestyle orientation may likely be predisposed to rely on private travel mode. The neighbourhood cluster on the north-west section of the region returned low coefficient values for demographic attributes, implying that the influence of demographic characteristics on mode choice is weaker in such neighbourhoods.

Socioeconomic characteristic variables show a weaker influence on travel mode choice in neighbourhoods where demographic factors show a stronger influence and stronger influence where demographic characteristic variables returned weaker influence. This pattern of relationship simply defines a spatial dichotomy that needs to be given detail attention in further study. However, the positive sign in the coefficient for factor 3 is an indication that such socioeconomic variables like highest education, household monthly income, number of cars in the household and number of driver's license in the household strongly affect the choice of travel mode in the north-west and north-east zones of the region. This means that in these zones affluent households that are well educated and have many cars with more members of the family having driver's license may prefer private mode of travel. This result is not unexpected as affluent families tend to depend on private cars since they can afford it and in most cases more members of the family may likely have a car to themselves. These neighbourhoods include Isiohor, Egor, Uwelu, Evbomore, Idunwowina, Iguosa, Ovbiogie, Ugbowo, Okhoro, Urubi, Uselu, Evbogida, Igue-Iheya, etc. These significant socioeconomic variables as evident from Table 3, increases the likelihood of choosing the private mode. From a spatial point of view, this is an expected result since these neighbourhoods clustered around high employment and education zone (Asikhia & Nkeki, 2013) and may be influenced largely by the presence of the University of Benin and the teaching hospital.

Gender which was extracted separately from other demographic variables in factor analysis (factor 5) also returned a significant coefficient with a negative sign in the local model. The neighbourhoods where the influence is higher are about the same as that of factor 2. The negative sign indicates that gender reduces the likelihood of choosing a private car. This indicates that the female folk in the region may prefer public transport mode over private car.

6.1. Planning and policy relevance

The implication of this research findings specifically, from a planning and public policy perspectives, is that there is spatial variations in the relationship between explanatory factors and peoples behaviour. Using these variations especially as it concerns the factors that may encourage people to choose a private car over public transport or otherwise would provide more realistic and detailed information for urban planning and policy formulation. For example, the local coefficient of GWLR predicted that as the socioeconomic status of the commuters living in the north-east cluster of neighbourhoods improves, the likelihood of driving a private car increases. This knowledge is a vital policy decision-making tool that would provide a platform for several policy options such as whether to engage a more drastic land use planning, reconfiguration and adjustment or to engage policies that would merely influence behaviour by encouraging private car owners to participate in car sharing/pooling. Policies like this would assist in achieving the fundamental objective of built environment and travel behaviour interaction studies which is hinged on the desire to reduce too much dependence on private car and in turn may lead to peak-hour vehicular traffic reduction, cut down exhaust emission and parking space demand.

In neighbourhoods where demographic attributes demonstrate stronger influence (i.e. south-east and south-west zones of the region) the policy option may be to provide modern effective mass transport system that is timely, comfortable, affordable, etc. This is because these zones are composed of older and larger size households with many jobs. These family compositions are major prerequisite for personal or household car(s) demand. Though it may be tough to make them drive less since they derive travel comfort in private car mode. The best practice is to deploy more comfortable mass transit system with latest technology bearing in mind the modern lifestyle orientation of these neighbourhoods commuters.

7. Conclusion

In conclusion, a spatial modelling approach was implemented here for the study of travel behaviour focusing on the interaction of built environment and at the same time, accounting for the influence of demographic and socioeconomic characteristics. The GWLR model proved to be most suitable for exploring this relationship since it accounted for local variation which is often lost when using global models such as GLR. The application of GWLR model proved to be valuable in analysing this kind of relationship in the study area because it was able to firstly, extend the modelling framework of GWR Gaussian model to accommodate binary categorical responses. Data on travel behaviour is often categorical in nature and this characterizes significant amount of researches emanating from applied geographical aspect of transportation and land use even though the spatial characteristics of the data used is not often accounted for. Secondly, geographical weighting of the survey data from BMR offered an in-depth explanatory tool that aided not only the exploration of global relationship but also accounted for the spatiality of the dataset and spatial heterogeneity in the relationship between travel behaviour and the covariates.

Using GLR model, it is expected that the relationship between travel behaviour and the explanatory variables are average out and a single value assigned to represent the parameter estimates for the entire neighbourhoods of the region therefore suggesting a uniform behaviour and relationships across the region. Implementing GWLR, parameter estimates for each variable is assigned to each neighbourhood and presented in a map. For example, the series of maps that were generated demonstrated the usefulness of the model's result in exploring the socio-spatial factors that may influence people's travel decision in BMR. The maps show that individual characteristics does affect travel mode choices, particularly the effect varies from neighbourhood to neighbourhoods. For example, GWLR model for the region was able to predict neighbourhoods that may depend on private car using the multi-dimensional set of predictors. This empirical evidence is substantive for major and specific urban transport policies since the result ascertained that travel mode choices varies over space and that socioeconomic and demographic attributes of people (which also varies across neighbourhood) significantly impact on this behaviour.

The explanatory variables for a local model of travel behaviour were first analysed using factor analysis which aggregated the variables into factor scores of built environment, demographic and socioeconomic factors. The model's estimation results suggest that built environment factors do not significantly count as a predictor of mode choice in BMR. Instead, this choice was shown to be locally associated with demographic and socioeconomic factors. Most importantly, the results indicated a spatial mismatch between demographic and socioeconomic characteristics, creating a dichotomy by demarcating the neighbourhoods into two levels of influence (demographic and socioeconomic). The implication of this is that effective location information is made available for planning allocation regarding transport investments, issues of excessive exhaust emission, energy use and congestion.

Conflicting with popular believe that built environment attributes are the key predictor that defines people travel behaviour, this study revealed that in a traditional African city people's individual or household characteristics are the key explaining factors of travel behaviour (which, in this study, is based on private and public modes). Several studies that came up with the former result were conducted in more advanced regions (such as North America and Europe). This perhaps points to the fact that regional variation in development may predetermine the outcome of such studies as this. The western region and other advanced regions of the world may have succeeded in reducing or perhaps significantly tackled the numerous issues inherent in the socioeconomic and demographic characteristics of their citizens but in Africa the case is largely different. The bane of most African cities is tied to corruption which has become a norm and has initiated endemic poverty. Consequent upon this, the people of this region have higher needs to first improve their socioeonomic and demographic status before concerning themselves with issues of built environment. So, it is not suprising that these individual characteristics would significantly influence their decision making processes. At this point, this study assert that the travel behaviour of people in less developed countries seems to be more influenced by their demographic and socioeconomic attributes while the travel behaviour of their counterpart in more developed countries seems to be strongly influenced by the built environment. This presents a note for further studies so as to approve or dispprove of this assertion.

Acknowledgements

We appreciate Digital Globe Foundation, Colorado, USA for the award of a high-resolution worldview satellite imagery of the study region as grant and support for this research. In the same manner we thank the two anonymous peer reviewers for their valuable and insightful comments on the previous version of the paper.

References

Abdel-Aty, M., Lee, J., Siddiqui, C., & Choi, K. (2013). Geographical unit based analysis in the context of transportation safety planning. *Transportation Research Part A, 49*, 62–75.

Aditjandra, P. T. (2013). The impact of urban development patterns on travel behaviour: Lessons learned from a British metropolitan region using macro-analysis in addressing the sustainability agenda. *Research in Transportation Business and Management, 7*, 69–80.

Aguero-Valverde, J. (2013). Multivariate spatial models of excess crash frequency at area level: Case of Costa Rica. *Accident Analysis & Prevention, 59*, 365–373.

Aguero-Valverde, J., & Jovanis, P. P. (2008). Analysis of road crash frequency with spatial models. *Transportation Research Record, 2061*, 55–63.

Antipova, A., Wang, F., & Wilmot, C. (2011). Urban land uses, socio-demographic attributes and commuting: A multilevel modeling approach. *Applied Geography, 31*, 1010–1018.

Asikhia, M. O., & Nkeki, F. N. (2013). Polycentric employment growth and the commuting behaviour in Benin Metropolitan Region, Nigeria. *Journal of Geography and Geology, 5*(2), 1–17.

Augustin, N. H., Mugglestone, M. A., & Buckland, S. T. (1996). An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology, 33*, 339–347.

Bahadure, S., & Kotharkar, R. (2015). Assessing sustainability of mixed use neighbourhoods through residents' travel behaviour and perception: The case of Nagpur, India. *Sustainability, 7*, 12164–12189.

Banister, D. (1997). Reducing the need to travel. *Environment and Planning B, 24*, 437–449.

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice methods and their applications to short-term travel decisions. In R. Hall (Ed.). *Handbook of transportation science* (pp. 5–34). Dordrecht, The Netherlands: Kluwer.

Bhat, C. R., & Koppelman, F. S. (1993). A conceptual framework of individual activity program generation. *Transportation Research, 27*, 433–446.

Bhatta, B.,B., Saraswati, S., & Bandyopadhyay, D. (2010). Urban sprawl measurement from remote sensing data. *Applied Geography, 30*, 731–740.

Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model. *Annals of Operations Research, 144*, 287–300.

Boarnet, M. G., & Crane, R. (2001). *Travel by design: The influence of urban form on travel.* New York: Oxford University Press.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis, 28*(4), 281–298.

Cao, J. (2014). Residential self-selection in the relationships between the built environment and travel behaviour: Introduction to the special issue. *The Journal of Transport and Land use, 7*(3), 1–3.

Cheng, L., Bi, X., Chen, X., & Li, L. (2013). Travel behaviour of the urban low-income in China: Case study of Huzhou City. *Procedia - Social and Behavioural Sciences, 96*, 231–242.

Comendador, J., López-Lambasb, M. E., & Monzónb, A. (2014). Urban built environment analysis: Evidence from a mobility survey in Madrid. *Procedia - Social and Behavioral Sciences, 160*, 362–371.

Connah, G. (1975). *The archaeology of Benin, excavations and other researches in and around*

Benin City, Nigeria. Oxford: Clarendon Press.
Crane, R., & Chatman, D. G. (2003). As jobs sprawl, whither the commuting? *Access, 23*, 14–19.
Deng, J. S., Wang, K., Hong, Y., & Qi, J. G. (2009). Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanisation. *Landscape and Urban Planning, 92*, 187–198.
Dieleman, F. M., Dijst, M., & Burghouwt, G. (2002). Urban form and travel behaviour: Micro-level household attributes and residential context. *Urban Studies, 39*(3), 507–527.
Du, H., & Mulley, C. (2012). Understanding spatial variations in the impact of accessibility on land value using geographically weighted regression. *The Journal of Transport and Land use, 5*(2), 46–59.
Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the American Planning Association, 73*(3), 265–294.
Fernandez, J. M., Chuvieco, E., & Koutsias, N. (2013). Modeling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. *Natural Hazards and Earth System Sciences, 13*, 311–327.
Flahaut, B. (2004). Impact of infrastructure and local environment on road unsafety logistic modeling with spatial autocorrelation. *Accident Analysis & Prevention, 36*, 1055–1066.
Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2002). *Geographically weighted regression: The analysis of spatially varying relationships.* Chichester: Wiley.
Fotheringham, A. S., Charlton, M., & Brunsdon, C. (1998). Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment & Planning A, 30*, 1905–1927.
Hadayeghi, A., Shalaby, A., & Persaud, B. (2010). Development of planning-level transportation safety models using full Bayesian semiparametric additive techniques. *Journal of Transportation Safety & Security, 2*, 45–68.
Haybatollahi, M., Czepkiewicz, M., Laatikainen, T., & Kytta, M. (2015). Neighbourhood preferences, active travel behaviour, and built environment: An exploratory study. *Transportation Research F, 29*, 57–69.
Huang, H., Song, B., Xu, P., Zeng, Q., Lee, J., & Abdel-Aty, M. (2016). Macro and micro models for zonal crash prediction with application in hot zones identification. *Journal of Transport Geography, 54*, 248–256.
Kim, C., & Wang, S. (2015). Empirical examination of neighborhood context of individual travel behaviours. *Applied Geography, 60*, 230–239.
Lee, J., Abdel-Aty, M., & Jiang, X. (2014). Development of zone system for macro-level traffic safety analysis. *Journal of Transport Geography, 38*, 13–21.
Lenth, R. V. (2011). *Java applets for power and sample size.* Retrieved Februray 1, 2016, from http://www.stat.uiowa.edu/~rlenth/Power.
Mathews, S. A., & Yang, T. C. (2012). Mapping the results of local statistics: Using geographically weighted regression. *Demographic Research, 26*(6), 151–166.
McGarigal, K., & Ene, E. (2014). *Fragstats: Spatial pattern analysis program for categorical maps.* Retrieved April 20, 2015, from www.umass.edu/landeco/research/fragstats/fragstats.html.
Meulman, J. J. (1992). The integration of multidimensional scaling and multivariate analysis with optimal transformations of the variables. *Psychometrika, 57*, 539–565.
Mitchell, R. B., & Rapkin, C. (1954). Urban traffic: A function of land use. *Journal of the American Planning Association, 55*, 24–36.
Nakaya, T. (2014). *GWR4 windows application for geographically weighted regression modeling. GWR4 user manual*.
Nakaya, T., Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine, 24*, 2695–2717.
Nkeki, F. N. (2016). Spatial-temporal analysis of land use transition and urban growth characterization in Benin metropolitan region, Nigeria. *Remote Sensing Applications: Society and Environment, 4*, 119–137.
Nkeki, F. N. (2018). *The influence of built environment, socioeconomic and demographic attributes on travel behaviour in Benin Metropolitan Region, Edo State, Nigeria*Unpublished Ph.D Thesis. Benin City, Edo State, Nigeria: Department of Geography and Regional Planning, University Benin.
Nkeki, F. N., & Osirike, A. B. (2013). GIS-based local spatial statistical model of cholera occurrence: Using geographically weighted regression. *Journal of Geographic Information System, 5*, 531–542.
NPC (2006). Population and housing census of the federal republic of Nigeria. *Priority Tables (LGA), 2* (National Population Commission).
Onokerhoraye, A. G. (1977). The spatial pattern of residential districts in Benin, Nigeria. *Urban Studies, 14*(3), 291–302.
Paez, A. (2006). Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. *Journal of Transport Geography, 14*, 167–176.
Pike, S., & Lubell, M. (2016). Geography and social networks in transportation mode choice. *Journal of Transport Geography, 57*, 184–193.
Pirdavani, A., Bellemans, T., Brijs, T., Kochan, B., & Wets, G. (2014). Assessing the road safety impacts of a teleworking policy by means of geographically weighted regression method. *Journal of Transport Geography, 39*, 96–110.
Pitombo, C. S., Kawamoto, E., & Sousa, A. J. (2011). An exploratory analysis of relationships between socio-economic, land use, activity participation variables and travel patterns. *Transport Policy, 18*, 347–357.
Rahul, T. M., & Verma, A. (2017). The influence of stratification by motor-vehicle ownership on the impact of built environment factors in Indian cities. *Journal of Transport Geography, 58*, 40–51.
Ramachandra, T. V., Aithal, B. H., & Sanna, D. D. (2012). Insights to urban dynamics through landscape spatial pattern analysis. *International Journal of Applied Earth Observation and Geoinformation, 18*, 329–343.
Reiff, B. (2003). *Statistical analysis of urban design variables and their use in travel demand models.* Oregon, USA: Oregon Department of Transportation.
Rodrigues, M., de la Riva, J., & Fotheringham, S. (2014). Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography, 48*, 52–63.
Sarvestani, M. S., Ibrahim, A. L., & Kanaroglou, P. (2011). Three decades of urban growth in the city of Shiraz, Iran: A remote sensing and geographic information systems application. *Cities, 28*, 320–329.
Selby, B., & Kockelman, K. M. (2013). Spatial prediction of traffic levels in unmeasured locations: Applications of universal kriging and geographically weighted regression. *Journal of Transport Geography, 29*, 24–32.
Strauss, J., & Miranda-Moreno, L. F. (2013). Spatial modeling of bicycle activity at signalised intersections. *The Journal of Transport and Land use, 6*(2), 47–58.
Sudhira, H. S., Ramachandra, T. V., Raj, K. S., & Jagadish, K. S. (2004). Urban sprawl: Metrics, dynamics and modelling using GIS. *International Journal of Applied Earth Observation and Geoinformation, 5*, 29–39.
Sun, B., Yan, H., & Zhang, T. (2017). Built environmental impacts on individual mode choice and BMI: Evidence from China. *Journal of Transport Geography, 63*, 11–21.
Taaffe, E., Gauthier, H. L., & O'Kelly, M. E. (1996). *Geography of transportation.* New Jersey: Prentice-Hall Inc.
Taubenbock, H., Wegmann, M., Roth, A., Mehi, H., & Dech, S. (2009). Urbanisation in India-spatio-temporal analysis using remote sensing data. *Computers, Environment and Urban Systems, 33*, 179–188.
Varotto, S. F., Glerum, A., Stathopoulos, A., Bierlaire, M., & Longo, G. (2017). Mitigating the impact of errors in travel time reporting on mode choice modeling. *Journal of Transport Geography, 62*, 236–246.
Vega, A., & Reynolds-Feighan, A. (2009). A methodological framework for the study of residential location and travel-to-work mode choice under central and suburban employment destination patterns. *Transportation Research Part A, 43*, 401–419.
Vovsha, P. (1997). The cross-nested logit model: Application to mode choice in the Tel-Aviv metropolitan area. *76th Annual meeting. Paper No. 970387*. Washington DC: Transportation Research Board.
Wang, C., & Chen, N. (2017). A geographically weighted regression approach to investigating the spatially varied built-environment effects on community opportunity. *Journal of Transport Geography, 62*, 136–147.
Wang, Y., Kockelman, K. M., & Wang, X. (2011). Anticipation of land use change through use of geographically weighted regression models for discrete response. *Transportation Research Record, 2245*, 111–123.
Wolf, L. J., Oshan, T. M., & Fotheringham, A. S. (2017). Single and multiscale models of process spatial heterogeneity. *Geographical Analysis, 50*(3), 223–246.
Xiong, Y., & Zhang, J. (2014). Applying a Life-oriented approach to evaluate the relationship between residential and travel behaviour and quality of life based on an exhaustive CHAID approach. *Procedia - Social and Behavioural Sciences, 138*, 649–659.
Zahabi, S. A., Miranda-Moreno, L. F., Patterson, Z., & Barla, P. (2012). Evaluating the effects of land use and strategies for parking and transit supply on mode choice of downtown commuters. *The Journal of Transport and Land use, 5*(2), 103–119.
Zhang, L., Hong, J., Nasri, A., & Shen, Q. (2012). How built environment affects travel behaviour: A comparative analysis of the connections between land use and vehicle miles traveled in US cities. *The Journal of Transport and Land use, 5*(3), 40–52.
Zhang, R., Yao, E., & Liu, Z. (2017). School travel mode choice in Beijing, China. *Journal of Transport Geography, 62*, 98–110.
Zhao, F., Chow, L., Li, M., & Liu, X. (2005). *A transit ridership model based on geographically weighted regression and service quality variables*Final report prepared for public transit office Florida: Florida Department of Transportation Lehman Center for Transportation Research. Retrieved April 4, 2013, from http://lctr.eng.fiu.edu/re-project-link/finaldo97591_bw.pdf.
Zwerts, E., Allaert, G., Janssens, D., Wets, G., & Witlox, F. (2010). How children view their travel behaviour: A case study from Flanders (Belgium). *Journal of Transport Geography, 18*(6), 702–710.