MODELLING THE RELATION BETWEEN URBAN MALARIA AND ENVIRONMENTAL INEQUALITY

A CASE STUDY OF CHILDREN, IN ABIA, NIGERIA

MSC Thesis

Chibunna Ann Orji

Student number: s6010172/4056965

Supervisor: Ir.P.W.M.Augustijn

Professor: Prof.Dr.M.J.Kraak

March,2017

SUMMARY

The risk of malaria falls heavily on under five aged children. Sub-Saharan Africa region bears the heaviest burden of malaria compared to other regions in the world (WHO,2015). This makes children from this region more vulnerable to malaria. Nigeria is one of the 6 countries in Sub-Saharan Africa that have about a quarter of the reported cases in the region. The record of the National Population Commission (2010) shows that malaria in the country, leads to 15 to 17% of fever cases in children and 300,000 child deaths per year.

Traditional malaria theories proof that children living in rural areas are more vulnerable to malaria than those living in urban areas. However, the studies of Austin (2014) & Fobil et al. (2014) are drawing emphasis on malaria epidemic in urban centres due to poor environmental conditions present in urban areas that could favour the breeding of mosquitoes. The exploratory analysis conducted in this research show children in the urban areas of the study area; Ugwunagbo and Aba South to have higher cases of malaria than those in the rural areas. This analysis also shows female children to have higher cases of malaria compared to male children.

The relation between **poverty** and **malaria** is a recurring finding in malaria research. The association between these two factors points to the relation between malaria and environmental inequality. This relation based on the report of WHOEurope (2010) has not been fully researched. This research therefore, uses the environmental inequality framework of Kruize et al. (2014) to define the relation between malaria and environmental inequality. From this framework, a model was derived - the multi-level differential malaria (MDM) model. This model groups the factors (vector based and host based factors) of urban malaria in 3 aggregation levels: macro level, ward level and the individual level.

Analysis was conducted on two levels (ward and individual level) to test the derived model. At the ward level; cluster analysis and Ordinary Least Square (OLS) regression were performed. The result of the cluster analysis identified four hot spot areas of malaria for the period of 2013 – 2015: Aba River, Aba Townhall, Igwebuike and Mosque. The OLS result showed houses that are within 1000 metres of the river and 200 metres of artificial water surfaces to be correlated with malaria.

 At the individual level analysis, data derived from the household survey; conducted during field work were used to conduct three regression analyses: Ordinal logistic regression, binary logistic regression and the OLS/spatial regression. The combined result of these analyses show: the use of preventive measures; room occupancy rate; income level and educational level to be strongly correlated with child malaria occurrence.

ACKNOWLEDGEMENT

TABLE OF CONTENT

1. INTRODUCTION

The effort to combat malaria epidemic and mortality has been successful over the last decade. The World Health Organisation(WHO) in 2015, reported a reduction of 37% in the number of malaria cases and 60% in the mortality rate among all age groups. Despite the improved records on malaria, over 200 million malaria cases and 400 thousand deaths were still reported.

Most of the malaria occurrence as reported by WHO, features in Sub-Saharan Africa. Nigeria is one of the 6 African countries that have about a quarter of the disease burden (Oduola, et al., 2013). The high occurrence of malaria in Nigeria is partly attributed to the topographic and climatic composition of the country. About 97% of the country is situated at low altitudes(National Population Commission, 2010). Furthermore, the climate is mostly wet and humid throughout the year, which provides a favourable breeding environment for mosquitoes (Umeh, 2013).

The main concern in the prevalence of malaria is its impact on children. Studies unanimously prove that children, especially those under the age of five bear the largest burden of malaria. In Nigeria, malaria has been assessed to be the main cause of about 15 to 57% of fever cases in children, resulting in about 300,000 child deaths per year (Abdulkadir et al., 2015 & National Population Commission, 2010). The highest occurrence of malaria in children is experienced during the first and second year of life, when immunity is still low (Umeh, 2013). The incidents of malaria in children are reported to be higher in rural areas than in urban areas (Esimai & Aluko, 2015). The lower transmission rate in urban areas as assessed by Kigozi et al.(2015) generally refers to fewer mosquito breeding sites, better access to treatment and a higher intervention coverage level.

Recent researches such as that of: Austin(2014) & Fobil et al.(2011) are drawing emphasis on malaria epidemic in urban centres due to the environmental conditions present in urban areas that could favour vector breeding and the rapid increase in urban population. Urban growth in sub-Saharan cities leads to social exclusion and creates an increasing level of poverty (Ngom & Siegmund, 2015). Poverty has been proven to have a strong relation with malaria. The association between poverty and malaria has been linked to economic and environmental aspects. Economic, due to lack of economic means to acquire clinical malaria control measures (Abdulkadir et al., 2015) and environmental, due to the poor dwelling conditions of the urban poor; conditions which are prone to mosquito proliferation (Yusuf, et al., 2010). This implies that children from poverty stricken households stand at greater risk in contacting malaria than others.

The burden of urban malaria in children from poor homes tends to create an inequality pattern in the distribution of malaria risk across the urban environment. This proofs the relation between malaria and environmental inequality. Environmental inequality as defined by Maantay (2002) is the disproportionate exposure of racial minority groups and the poor to environmental hazard and its consequent effect on health and the environment. Though environmental inequality is widely studied in pollution studies, according to Evans &

Kantrowitz, 2002)Evans & Kantrowitz (2002) inequality also includes other environmental hazards such as crowding, water quality and poor neighbourhood conditions which are strong influences of malaria.

The association between malaria disease and environmental inequality has not been extensively researched. The report of World Health Organisation Europe(2010) show that the link between environmental factors and malaria has not been established. This could explain the clinical and preventive approach adopted in many malaria studies. It is increasingly becoming evident based on recent studies such as that of De Silva & Marshall (2012), that certain environmental factors play a profound role in malaria transmission. This recognition has set the framework for causative studies on malaria epidemic.

Linking environmental inequality and urban malaria requires an understanding of the constituting factors that influence the two key aspects of inequality and these include: exposure and health outcome. These aspects are synergistically related, it is thus apparent, that exposure levels to environmental risks are likely to portend adverse health consequences (Deguen & Zmirou-Navier, 2010). However, health outcomes are not solely dependent on exposure, but on other factors external to exposure such as personal and behavioural factors. Malaria is therefore the health outcome of the exposure to certain environmental risk factors and individual factors. Children by nature are victims of malaria due to low immunity. This explains the individual/personal risk factor. Vulnerability increases when children are exposed to malaria through their living environment. Studies of Abdulkadir et al (2015) & Yé et al (2006) prove that children from poor homes have greater exposure to malaria than those from other income level. This, in essence, validates the notion of the interlinked relationship between environmental inequality and malaria prevalence.

1.1. Research objective
The aim of this study is to use existing frameworks in environmental inequality to define the relationship between inequality and malaria. The study further tests the framework on a case study for children under the age of 5 in Ugwunagbo and Aba South Local government areas, in Abia, Nigeria.

The research splits into 2 objectives and sub-objectives.

The first research objective is to (derive a framework from existing inequality frameworks that models the relation between the influencing factors of environmental inequality and urban malaria) investigate the environmental inequality factors in existing frameworks that are associated with urban malaria.

- To investigate the environmental, social-economic and demographic factors that influences the occurrence of malaria in urban areas.
- To investigate the existing frameworks in environmental inequality
- To describe the factors that influence environmental inequality from the framework

- To align the factors of environmental inequality and urban malaria to determine their association

The second objective is to analyse the impact of environmental inequality on the spatial and temporal pattern of malaria occurrence in 0-5 aged children of Ugwunagbo and Aba South in Nigeria?

- To identify hot and cold spot cluster locations of malaria occurrence in under 5 children over time
- To identify the spatial inequality factors that characterizes both hot spot and cold spot locations
- To analyse the spatial inequality factors that correlate with the malaria occurrence in the varying locations.

1.2. Research framework

This research undertook 3 steps as shown in figure 1.1 to realize its objectives. These steps include: literature based review, linking of concepts and analysis.

The literature review in Chapter 2 reports the factors that cause urban malaria and those that cause environmental inequality. The factors are explained in two groups: Vector-based factors (factors that enable the breed of malaria vector) and host-based factors (factors that enables the transmission of Plasmodium virus from vector to hosts; humans). On environmental inequality, the research uses the multilevel model as described by Kruize et al. (2014) to explain the various factors that cause inequality. This model consists of 3 levels which include: macro or Institutional, community level and individual level.

The second step links the findings of the two concepts (environmental inequality and urban malaria) to derive the relation that exists between them. The derived link is an adjusted model "multi-differential model" based on the multilevel model of Kruize et al (2014) and the 3 differential levels of environmental inequality (Deguen & Zmirou-Navier, 2010; Maantay, 2002) which include: differences in the distribution of hazards, differences in exposure and differences in health outcomes. This is discussed at the end of chapter 2.

Chapter 3 discusses the study area and the data. It further presents the exploratory analysis done on the malaria data.
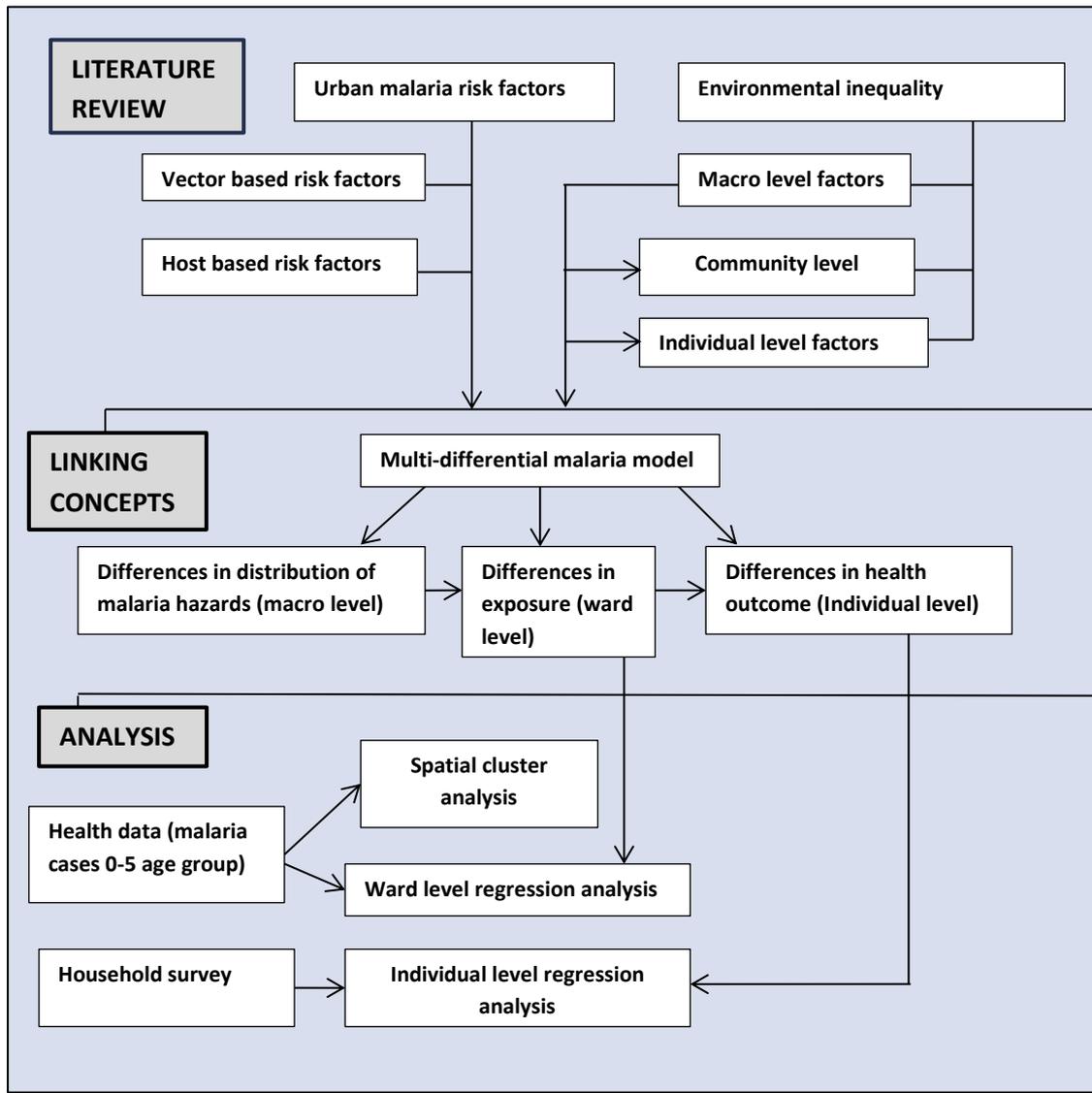
Figure 1:1: Conceptual framework

In the final step (analysis), 3 analyses are performed: Cluster, Ward level regression and Individual level regression analyses. The last two analyses on regression, tests the derived multi-differential malaria model. The method applied for these analyses is discussed in Chapter 4 while analytical results are discussed in Chapter 5. Chapter 6 gives the conclusion, reflection and recommendation.

2. URBAN MALARIA AND ENVIRONMENTAL INEQUALITY

This chapter gives a literature based review on malaria and environmental inequality. It therefore splits into two sections. The first section discusses malaria species and transmission, the concept of urban malaria and the risk factors associated with urban malaria and their implication on children. The second section discusses environmental inequality and the factors that cause inequality. The inequality causal factors are explained using the multilevel framework as described by Kruize et al. (2014).

2.1.   Malaria species and transmission

Malaria is transmitted by the female anopheles mosquito of the genus plasmodium (Martens et al., 1999). Mosquitoes depend on blood meals of their host (humans or, animals) to produce eggs and to extend its life span. The mosquito vector is either anthropophilic; dependent on humans for blood meals or zoophilic; dependent on animals for blood meals (Rosas-Aguirre et al., 2015).

According to Martens et al (1999) there are four species of the Plasmodium malaria parasite transmitted by the female anopheles vector and they include:

- Plasmodium vivax which has the most extensive geographic range and is present in many temperate zones as well as the tropic and sub-tropic zones.
- Plasmodium falciparum is the most common species in tropical areas and the most dangerous clinically.
- Plasmodium ovale resembles vivax and it is found mostly in West Africa.
- Plasmodium malariae is less apparent than the other species and is found mainly in tropical Africa.

Plasmodium falciparum, Plasmodium ovale and Plasmodium malariae are the dominant species of malaria parasite in Nigeria. Of these 3 species, Plasmodium falciparum is found to be the most prevalent and the major cause of malaria disease. The Nigerian malaria survey conducted in 2010 showed that 95% of the surveyed children were infected with Plasmodium falciparum. The other surveyed children were diagnosed with Plasmodium ovale and malariae.

There are 422 anopheles vectors distributed around the world, however, only about 40 species transmit malaria in the genus of Plasmodium, the main virus causing malaria (Martens et al., 1999). The dominant anopheles vectors in Nigeria are the Anopheles gambiae, Anopheles funestus and the Anopheles arabiensis. These anopheles species are strongly anthropophilic. The anopheles Gambia is the highest transmitter of Plasmodium falciparum. They are responsible for the highest number of bites per person per year when compared to the Anopheles funestus, affecting mostly under 5 aged children (National Population Commission, 2010).

The anopheles gambiae are of 2 forms: the S and the M molecular form (De Silva & Marshall, 2012). The S molecular form of the vector adapts better in rural and humid forest

areas while the M molecular form adapts in urban areas. The findings of Oduola et al. (2013) show the presence of both forms of anopheles gambiae in all the urban communities sampled in Oyo State, Nigeria. The gambiae anopheles vector was previously understood to have a strong preference for and only adapts to clean water (Afrane et al., 2012). However, the findings of Marquetti et al. (2007) conducted in Jigawa State, in Nigeria show the adaptation of the anopheles gambiae to polluted water surfaces in the urban areas.

2.2. Urban malaria

It has been proven in several researches such as that of: Esimai & Aluko (2015) ; Kigozi et al. (2015) & Robert et al. (2003) that urban areas have a lower malaria incidence than rural and peri-urban areas. Low transmission of malaria in the urban environment has been linked to better access to health facilities and the negative effect of pollution on the growth of malaria vector (Fobil et al., 2011). Urban areas also have fewer breeding sites for the malaria vector such as natural water surfaces (Kigozi et al., 2015).

There are contrasting studies focusing on the favourable factors that could increase the vulnerability of urban areas with malaria. Fobil et al. (2011) explains that urban areas create environmental conditions that favour insect vector breeding. The urban environment has several artificial breeding sites that provide protected habitat for all life stages of the malaria vector. Breeding sites come in the form of water surfaces such as drains, canal, pools, domestic water reservoirs and hollow surfaces such as foundations of unconstructed buildings and tyre tracks on road surfaces (De Silva & Marshall, 2012). The studies of De Silva & Marshall(2012) & Ogunrinola & Adepegba(2012) also counters the notion that pollution inhibits the growth of the malaria vector. These studies show that the Anopheles gambiae vector; one of the deadly transmitters of the malaria parasite, adapts to polluted water in the urban environment. The vector was found in water sources with high concentration of heavy metals such as iron, copper and lead, in waste disposal sites and septic tanks having contaminated human faeces.

The prevalence of malaria in urban areas has also been viewed in relation to social factors such as poverty. Urban growth in most African cities leads to social exclusion that creates increasing levels of poverty. This has to do with the fact that growth in urban centres occurs without industrial development (Ngom & Siegmund, 2015). Cities still possess some rural characteristics such as urban farming; poor housing quality and high room occupancy rates. Housing quality illustrates the social status of households and it is the case that the urban poor dwell in houses with poor housing infrastructure and in slum dwellings. This has placed children from poor households at greater risk of malaria because these poor housing features facilitate malaria transmission rates.

The analysis of Ngom & Siegmund (2015) & Yé et al. (2006) revealed that the poorer sections of the population had greater chances of having malaria. Their findings show that occupants with high quality housing are 50% less likely of having malaria than occupants living in poor housing conditions. Poor households are less able to afford prevention or treatment and the higher burden of malaria may push these same individuals deeper into poverty (Rosas-Aguirre et al., 2015).

Malaria poses a serious challenge in urban areas not for the growing number of factors that favour the transmission of the disease, but for the spatial heterogeneity that characterizes malaria transmission. Robert et al (2003) asserts that there are considerable variations in the level of transmission that exist amongst cities and within different districts in the same city. Variations in malaria occurrence are strongly dependent on the environmental and living conditions in urban areas (Fobil et al., 2011).

The variations in urban malaria show inequality patterns in the transmission levels between areas and groups. The transmission level of malaria in urban areas might be low when compared with peri-urban areas and rural areas, but the tendency of concentration of malaria transmission within a vulnerable locality and age group, children, is worth investigating.

2.3. Urban malaria risk factors

Malaria is enhanced by certain factors which either increases the chance of vector breeding or vector to host transmission. These factors are referred to as risk or causal malaria factors. Risk factors of malaria in urban areas are complex in nature and were not fully understood in the last decades (Donnelly et al., 2005). The complex nature of urban risk factors is based on the fact that the set of factors generally understood as risk are more of natural factors and often do not apply to urban areas. De Silva & Marshall (2012) explains, that natural risk factors such as vegetation, clean water surfaces such as lakes, rivers and swamps, known breeding malaria sites are fewer in urban areas and more in rural areas. This explains the heavy burden of malaria in rural areas compared to urban areas. Though some of these natural factors are present in urban areas, a greater percentage of risk factors are artificial and man-made.

Present malaria studies such as De Silva & Marshall (2012) & Ngom & Siegmund (2015) have attempted to identify artificial risk factors inherent in the urban environment. These studies revealed that the risk factors of urban malaria do not only lie on factors that favour the growth and existence of the vector species, but on a wide range of social and environmental factors that enables vector to host contact.

The subsequent sections make a comprehensive review of the risk factors in urban areas as identified in literature. Discussion is made in two groups: vector-based urban risk factor and host-based factor.

2.3.1. Vector based risk factors

These are factors that favour breeding of the malaria vector. Factors are both natural (climatic and topographic) and artificial or man-made factors.

2.3.1.1.   Climatic factors

Rainfall and temperature are the climatic factors that affect the breeding of the malaria vector. Both factors are needed at suitable conditions to increase the breeding of mosquitoes (Tanser et al., 2003).

Rainfall

The abundance of anopheles mosquitos is strongly affected by rainfall events. A high precipitation increases water surfaces that harbour malaria vector. Mosquitoes breed in standing waters such as freshwater pools or marshes (Martens et al., 1999) and also in temporal and polluted water surface (De Silva & Marshall, 2012). The capacity of these water surfaces is saturated by the outpour of rain. Consequently, this results in a larger pool of surface water that favours breeding of the vector and the longevity of the adult mosquito.

The study of Afrane et al. (2012) proves a negative correlation between rainfall and adult and larval abundance. They explain that higher amounts of rainfall wash larvae out of its habitat and thus, reduce the abundance of mosquito larvae. This consequently, leads to a reduced number of positive larval habitats during rainy season.

Temperature

Temperature has a direct effect on the duration of the sporogonic (larvae) life cycle of the malaria parasite and vector survival (Tanser et al., 2003). At a minimum temperature reaching freezing level, the population of anopheles vector radically reduced. Though it has been proven by the study of Petersen et al. (2013) that malaria vector can develop in temperate zones even in zones with freezing temperature. The plasmodium vivax was dominant in the temperate zones of the world before the 1950's and presently in temperate zones of Italy and South Korea. Malaria in temperate zones is as a result of the long incubation period of the genotype of plasmodium vivax and the migration of people from malaria endemic regions.

Vector density increases at a warmer temperature of 19.5°C and above. A consistent temperature of 19.5°C prolongs the duration of the sporogonic cycle of the plasmodium falciparum parasite and guarantees- 4% of the total vector cohort surviving (Tanser et al., 2003). Temperature of over 32–34°C rapidly reduces the survival rate of parasite.

The minimum temperature for both parasite development and transmission lies between 14.5 and 15°C in the case of Plasmodium vivax and between 16 and 19°C for Plasmodium falciparum (Martens et al., 1999; Petersen et al., 2013).

It is reviewed by studies such as Craig et al. (1999) & Tanser et al. (2003) that climatic factors (rainfall and temperature) have serious impact on malaria occurrence. The results of the temporal analysis conducted by these studies show that climatic condition in Nigeria for almost all the monthly period (7-12 months) of the year are suitable for Plasmodium falciparum transmission. Despite the impact of climatic conditions in malaria transmission, the studies of Ngom & Siegmund (2015) &Robert et al (2003) emphasize the significant impact of socio-demographic and socio-environmental factors on urban malaria. These factors are discussed in section 2.3.2.

2.3.1.2. Topographic factors

Altitude

Altitude is generally considered to play an important role in limiting malaria in the tropical highlands by negatively influencing the breed of vector species (De Silva & Marshall, 2012). Areas at a high altitude have low transmission when compared to areas at low altitude. The increased water run-off downstream and warmer temperature in lowlands makes low altitude areas a preferred breeding location for malaria parasite.

However, the studies of Ernst et al.(2006) & Peterson et al.(2009) show the possibility of malaria transmission in highland areas. These studies confirm a clustered pattern of malaria in the highlands of Ethiopia and Kenya with altitudes above 1600m and 2000 meters. Areas that were prone to malaria risk in highlands are lower altitude areas within high plains. These include valleys and lower depressed areas. Vector density could form clusters in these depressed areas causing risk to the inhabitants living at close distance to these sites.

Urban agriculture

Urban agriculture has been used as a measure to enhance food security and to alleviate poverty in urban areas. Besides these foreseen benefits, they have been found optimal sites for vector breeding (Klinkenberg et al., 2008).

Irrigated agricultural fields are the most attractive habitats for mosquitoes. Dug out wells present in these fields, foot prints, furrows and seepages contain the largest amount of mosquito larvae which consequently increases the vector population (Afrane et al., 2012; Klinkenberg et al., 2008). Hamilton et al. (2013) asserts that the proportion of habitats containing Anopheline larvae and adult anopheles is 1.7 times greater in urban areas with agriculture than those without agricultural fields.

The abundance of larvae and adult mosquito in urban agricultural fields has serious implication on those who live and work in urban agricultural areas, irrespective of other factors such as urban or peri-urban location (De Silva & Marshall, 2012). Children living in close proximity to agricultural areas, especially irrigated fields, are at high risk of contracting malaria. The study of (Klinkenberg et al., 2006) conducted in the urban area of Accra, Ghana show a significant malaria epidemic on children living close to urban agriculture.

Animal husbandry is another aspect of agriculture that increases the presence of the malaria vector. Animals increase the attraction of zoophilic mosquitoes under the stimulus of CO_2 and octanol. When mosquitoes are both zoophilic and anthropophlic such as the Anopheles Albimanus, children living or playing close to livestock farms becomes vulnerable to malaria (Rosas-Aguirre et al., 2015).

Natural water surfaces

Natural water surfaces such as rivers and floodplains provide great breeding grounds for mosquitoes in riverside urban communities (De Silva & Marshall, 2012). River banks are rich in aquatic plants and these plants retain sunlight, resulting in lower temperature conducive for vector breeding. However, vectors could be reduced by the presence of predators such as fish species, tadpoles and aquatic insects of the order coleopteran and hemipteran species (Akono et al., 2015).

2.3.1.3. *Artificial/made-made factors*

Urban areas have several artificial sites that favour the breeding of the mosquito vector. These sites are in the form of artificial surface water in the surrounding of residential communities and waste dumpsites.

Artificial surface water as pointed by De Silva & Marshall (2012) & Klinkenberg, et al (2008) include: blocked and stagnant drainage channels, road surfaces with tyre tracks and potholes, stagnant swimming pools, domestic water reservoir, broken water pipes and pools at construction sites.

Blocked drains are often due to poor sanitation which leads to a reduced water flow. Consequently, stagnant water pools accumulate, forming suitable sites for the breed of mosquitoes.

Unsurfaced roads and poorly maintained road network raise the risk of tyre tracks and potholes. They create shallow surfaces having temporal pools of water that are saturated during rainfall. Tyre tracks are more common in areas of high socioeconomic status, which tend to house more vehicle owners having poor conditioned roads.

Poor waste management has serious implications on malaria transmission. Solid waste in Nigeria and in most Sub-African countries is generated in proportionately large amounts without adequate management set in place. A large proportion of solid waste is dumped either in poorly managed landfills or in open dumps and this constitutes a source of health risk to surrounding residents (Ogunrinola & Adepegba, 2012).

Waste sites provide a favourable breeding site for the malaria vector due to the dampness of biodegradable waste and the contaminated effluence that is discharged, especially during the downpour of rain (Wachukwu & Eleanya, 2007).

It has been proven that people living or working at close distance to waste disposal sites are of greater risk of contracting the plasmodium virus than those living further away. The study of Afon (2012) reports a significantly higher level of malaria parasitaemia on scavengers (informal collectors of valuable waste from dump sites) working in the waste dump sites. Similar findings were made by Wachukwu & Eleanya (2007) in Lagos where malaria prevalence was found to be higher amongst regular on-site workers than off-site or unexposed workers of the open waste dump site in Lagos. The study of Rahman (2006) identified waste as one of the strongly associated factors to the prevalence of malaria in

Aligarh city of India. The implication of these findings on children is that infants attending schools close to waste dump sites or living and playing close to these sites could be prone to malaria infection.

2.3.2. Host based risk factors

Host based factors are those socio-environmental, socio-demographic and metric factors that enables malaria transmission from malaria vector to their host, humans. They have been assessed by Ngom & Siegmund (2015) & Robert et al. (2003) as key influences to malaria epidemics and have a larger impact than the vector based factors previously discussed. The prevalence of malaria is strongly based on host factors because without them, there is no malaria transmission and no account of malaria epidemic or incidence.

2.3.2.1. Socio environmental risk factors

Housing quality and condition

The quality of housing is one of the socio-environmental risk factors of malaria. Quality is assessed by the type of material used in housing construction which could be durable or natural. Houses built with durable materials comprise of: brick, cement, tile for walls and asbestos or metals for roofs. Houses built with natural building materials comprise of thatched or mud walls and thatched or other plant materials for roofs (Leandro-Reguillo et al., 2015).

The findings of Konradsen et al. (2003) & Lwetoijera et al. (2013) show that houses built with durable materials have better resistance to the entrance of mosquitoes indoors than houses built with natural materials. These authors explain that houses made of mud walls and grass roofs often have crevices used by mosquitoes to enter the house. Also of major importance is the cool and dark condition that characterizes mud wall and grass roof houses which provides hiding and resting place for mosquitoes.

Housing quality has also been viewed in relation with condition of building materials (Konradsen et al., 2003; Lwetoijera et al., 2013). Houses built with durable materials can deteriorate in quality over time due to poor maintenance. This is seen by cracks in walls, opening or gaps in the eaves and lack or damaged screening over windows. Houses in these deplorable conditions provide access points for mosquitoes to enter indoors (Lwetoijera et al., 2013). The study of (Yé et al., 2006) shows that children living in houses with natural roofing materials and open eaves had a higher malaria infection than children living in suitable roofed houses. These children are 30% more at risk of malaria transmission than their counterpart (Konradsen et al., 2003).

2.3.2.2. Socio-demographic risk factors
Income level

Income level plays a profound role in the transmission of malaria in children. Infants from low income or poor families are more susceptible to the risk of contracting malaria than those from average and wealthy families. It is clear from findings such as (Nriagu et al.,

2008) that low income families are more likely to live in overcrowded areas with open and contaminated sewers. They also tend to reside in poorly constructed buildings which are porous for the entry of mosquitoes indoors. Poor households are faced with inadequate water supply and poor sanitation which increases vector breeding and contact. Also from an economic viewpoint, low income families have a lower capacity of using health care services and preventive malaria control measures (Robert et al., 2003).

Educational level

Lower education level of parents and caregivers contribute to malaria transmission in infants. Possible reasons for this is that uneducated parents tend to respond late to malaria symptoms and they lack awareness of malaria control measures used in the protection against malaria (Abdulkadir et al., 2015; Robert et al., 2003). Also, children having low educated parents might not attend schools and hence tend to play more often outside. This increases exposure levels of infants to vector bites, especially when the surrounding environment is prone to vector proliferations (Nriagu et al., 2008).

Household size and density

This factor has been identified as a significant demographic factor in the transmission of malaria. Houses with more occupants tend to attract more mosquitoes than households with fewer occupants. The relation between household size and malaria is merely not based on the numeric composition of a household (number of occupants) but on spatial composition (the number of occupants per room space). Ngom & Siegmund (2015) explains that a high sleeping density room favours the emanation of human odour that attracts the malaria vector. Mosquitoes identify and find their host through olfaction. Substances such as lactic acid and ammonia present in the human skin and carbon dioxide exhaled through breathing attract mosquitoes to their host (Rosas-Aguirre et al., 2015).

2.3.2.3. Metric risk factors
Distance to vector breeding site

The presence of vector breeding sites is of no effect in malaria transmission except when such sites are in close distance to the location of host. Distance plays an important role in the vector-host malaria transmission and it has been found a reoccurring risk factor in all malaria studies. There is a consensus in all malaria studies that people living in close proximity to vector breeding site are at higher risk of contracting malaria parasites than those living further away. However, variations vary in the exact distance in meters that poses serious risk to malaria transmission. DeSilva & Marshall(2012) and Ogunrinola & Adepegba(2012) specify a threshold distance of 250 meters to vector sites as hazards prone areas. Konradsen et al (2003) in his analysis sets a threshold distance of 750 meters.

Distance to health centre

Access to health centre especially for children under the age of 5 has the potential of reducing significantly the number of malaria cases and death(Rutherford et al., 2010). This with reason that they have access to effective health care services such as: treated bed nets(Larson et al., 2012). The study of Feikin et al. (2009) show that for every 1 km increase(up to 4km) in the distance from health centre, there was a corresponding increase in the number of malaria cases. Although distance to health centre plays a role in the reduction of malaria, there are also barriers that could limit its impact. Barriers such as: educational, cultural and financial factors(Kizito et al., 2012).

Migration pattern

The prevalence of malaria in urban areas can be increased by two types of movement: rural to urban migration; this is often a permanent relocation to urban areas for purposes of jobs and studies and urban to rural migration; this is often temporal and is referred to as rural travel (Robert et al., 2003). In both cases, migrants bring infection from rural endemic areas to urban areas where they reside. Children involved in rural travel for vacation for instance, could be at risk of exposure to mosquitoes. The result of Klinkenberg et al. (2008) conducted in Ghana show that children that travelled to rural areas within the 3 weeks of sample period had a higher chance of malaria parasite than those who did not embark on rural travel.

2.3.2.4. Clinical factors

IRS (Indoor residual spray) and ITN (Insecticide treated nets) are the clinical control measures used in the prevention of malaria. The lack of use or ineffective use of these measures could lead to malaria (Esimai & Aluko, 2015).

The study of Lwetoijera et al. (2013) shows that these control measures are not effective in reducing malaria transmission, especially in poor quality houses with open eaves and roofs. Their study shows a high number of malaria vectors indoors in poor conditioned houses despite the use of IRS and ITN control measures. Ineffectiveness of IRS and ITN is also linked to its lack of protection from outdoor mosquito bites since they are only used indoors. The implication of this on children is that clinical measures do not protect children playing close to vector prone areas and even when measures are used indoors, children living in poor quality homes are still at risk.

2.3.2.5. Individual factors

The vulnerability to malaria has been linked to biological components such as immunity. Infants generally have a low level of immunity, especially during the first and second year of life (Umeh, 2013). This makes them highly susceptible to malaria infection. Besides low immunity levels, factors such as nutrient deficiency and existing illness might decrease considerably, the immune response of children to malaria (Deguen & Zmirou-Navier, 2010).

The study of (Nriagu et al., 2008) raises counter claim to immune deficiency and existing illness. Their study shows that a high blood level of lead provides immunity against malaria

infection in children. Though presumptively, lead may affect the effectiveness of malaria vaccines due to the changes lead poisoning has on the immune regulatory function.

Summary

Malaria as has been discussed is transmitted by the female anopheles mosquito. The most dominant malaria specie in Nigeria is the Plasmodium falciparum, transmitted by the vector, Anopheles gambiae. This class of specie and vector is one of the deadliest transmitters of malaria. They are responsible for about 95% of bites in Nigerian children. Urban children have a relatively lower rate of malaria transmission than those in peri-urban and rural areas. The attention of malaria in urban children is on the key notion of the possibility of cluster formation of malaria in certain locations owing to poverty related factors. The risk factors associated with malaria are summarized in Table 2.1. It is worthy to note that host-based factors are key factors in malaria transmission where most factors suggest inequality.

Group	Urban Risk factors
Vector-based	Climatic • Temperature • Rainfall
	Topographic • Altitude • Natural water surfaces • Urban agriculture
	Artificial • Man-made water surfaces • Waste
Host-based	Socio-environmental • Housing quality(wall, roof, screening facilities) • Housing condition • Housing location
	Socio-demographic: • Occupancy rate • Educational status • Income level
	Metric: • Distance to vector site • Distance to health facilities • Migration pattern (urban to rural and rural to urban travel patterns).
	Clinical: • Medical control measures Individual: • Immunity(Dietary and • Existing illness

Table 2:1: Risk factors of urban malaria.

2.4.   Environmental inequality

This has been defined from different standpoints. Banerjee (2013) explains these standpoints as social, political, science and technological and environmental.

From a social perspective, environmental inequality is viewed as the unequal distribution of environmental and economic resources in a society, with distributive pattern predicted by race and income or social class (Maantay, 2002; Pellow, 2000).

Political researchers see environmental inequality as the segregation of the minority from political freedom, social choices, political right and institutional participation in the environmental decision making process.

From a science and technological standpoint, environmental inequality is viewed as the suppression of local/community knowledge claim by expert knowledge on environmental issues.

Environmental inequality from an environmental perspective is the outcome of both social and political processes. It is defined as the uneven spaces brought about by distributional inequity and inequality in the implementation of environmental policies which, consequently, leads to the vulnerability of marginalized communities to environmental risk.

The social, political and environmental viewpoints have formed the basic framework in environmental inequality studies. This framework is built on the relation between environmental quality and social hierarchies (Pellow, 2000). O'Neill et al. (2003) explains, that at each step of the social hierarchy, individuals tend to have a better qualitative environment and a better health compared to those immediately below them.

 Maantay (2002) gives a comprehensive definition of environmental inequality, addressing both political and environmental inequality aspects. Environmental urban inequality based on Maantay's definition is the "*disproportionate exposure of communities of colour and the poor to environmental hazard and its consequent effect on health and the environment, as well as the unequal environmental protection and environmental quality provided through laws, regulations, programs, enforcement and policies*".

According to Brulle & Pellow (2006), the vulnerable group exposed to environmental hazard are not necessarily the racial/ethnic or poor income group, although, they bear the highest environmental risk and suffer worst health outcomes. Effects extend to ranges of middle-classed income group and to other vulnerable social group such as: the young and old, women, immune compromised and the infirmed (Maantay, 2002; O'Neill et al., 2003).

The central notion in the discussion of environmental inequality is that certain locations and social groups, mainly the poor and ethnic minority bear the heaviest burden of environmental risk than others. Such impact is influenced by 3 levels of differences:

- *Differences in the distribution of environmental hazards:* Environmental hazards refer to those noxious/toxic facilities that emit pollutants which adversely have a negative impact on health and wellbeing (Pellow, 2000). The hazard often emphasised in

environmental equity studies are anthropogenic (pollutant) facilities. However, hazards based on the description of WHO Europe (2010), are also meant to include: poor housing, poor neighbourhood conditions, crowding conditions and poor water quality.

Studies in environmental equity proof that environmental hazards are unevenly distributed across the urban space. Hazards tend to form clusters in locations occupied by socially disadvantaged groups. The reason for the heterogeneity in the spatial pattern of hazards has been attributed to the systemic racism and class preferences in landuse decisions which portions hazardous facilities away from wealthy neighbourhoods to poorer ones (ONeill et al., 2003). The studies of Banerjee (2013) & Pellow (2000) make counter claims to the notion of intentional discriminatory practices in hazard distribution. Distribution of inequality as explained by these authors is based on social and economic processes of the state which inevitably affect the deprived social groups. These processes are discussed in details in section 2.5.3.

- *Differences in exposure:* Inequity in the spatial distribution of environmental hazard invariably creates differences in exposure levels. The victimized population residing in the same boundary close to an environmental hazard face higher exposure to environmental risk than others. This is evident in residential segregated communities where population of similar colour and socioeconomic class are bound by the same post code, census tract and county boundaries. Besides residential exposure, differentials in exposure risk has also been linked to occupational (work) and lifestyle(behavioural) factors (Maantay, 2002). Victims of exposure are often the low-waged, low income social groups, who are exposed to hazardous working environment and reside in environmental risk prone areas (O'Neill et al., 2003).

- Differences in health outcomes: Researches refer to this level of difference as environmental health inequality. A recurring evident established in health inequality findings is that low social groups bear the heaviest health effects of environmental risk than others, portrayed in differing diseases and ailment. Health impact is not solely dependent on the level of exposure to hazard, but on vulnerability and susceptibility factors (Deguen & Zmirou-Navier, 2010 & WHOEurope, 2010).
These factors include: health status, biological sensitivity and immunity, age, gender, poverty and access to better medical treatment. Deguen & Zmirou-Navier(2010) explains that even in the case of low exposure levels; these vulnerability factors play a significant role in increasing health effects of hazard among social disadvantaged groups. At high risk exposure, a synergistic effect of both exposure and vulnerability could produce even increased health consequences.

2.5. Causal factors of environmental inequality

Previous studies explained the causes of environmental inequality as an outcome of the discriminatory practices of the society where the poor are intentionally placed or exposed to environmental hazards because they are less powerful than the state (Banerjee, 2013). Pellow (2000) explains this as the perpetrator-victim model. The study of Brulle and Pellow (2006) and Banjee (2013) criticizes this model and stresses emphatically, that environmental inequality emerges through a process of ongoing societal changes. It is explained, that as a society undergoes certain transformation, certain factors interact which widens the gap between social hierarchies. This consequently increases the exposure of low hierarchy groups to environmental risk (Brulle & Pellow, 2006). These factors are a complex set of social, economic, political, psychological and environmental factors (Kruize, 2007). They have an interlinked relationship and partially overlap which makes it difficult to concisely define their effects on inequality.

Several theories and models (Pellow, 2002; Brulle and Pellow, 2006 and Kruize, 2007) have attempted to describe the contribution of these factors to environmental inequality. However, according to Kruize (2007) these models are not coherent and they tend to partially describe the factors influencing inequality. Thus, none of the theories or models gives an all-inclusive description of the various factors that causes environmental inequality.

Recent frameworks such as that described by Kruize et al. (2014) have attempted to provide an integrated model that helps explain the causes of environmental inequality and the interdependencies that exists between them. This model has been represented by Kruize et al. (2014) using a multilevel model. It consists of 3 interlinked levels and they include: macro or societal level, community level and the micro or individual level. Figure 2.1 shows these levels and their constituting factors.

2.5.1. Micro/individual level

This level refers to factors that pertain to an individual. According to O'Neill et al. (2003) and Soobader et al. (2006), they include: demographic (age, gender, race, income and educational status); behavioural (health behaviour) and acquired factors (dietary and nutritional status, proximity to health services and pollutants). These factors create differences in vulnerability to the health impact of environmental risk, that even given the same environmental condition; they exert a significant level of influence on health outcomes (Morello-Frosch & Lopez, 2006).

O'Neill et al. (2003) explains that children and the aged population, non-indigenous population, and population group with low income and educational level are the groups vulnerable to the health impact of environmental risk. The same applies to Individuals with poor health behaviours, poor diet and nutrition, lack of access to health care and those living or working close to hazardous facilities.

17

According to Morello-Frosch & Lopez (2006), the micro-level factors influence vulnerability indicators such as: existing illness and exposure-response (Immunity) function. This consequently leads to negative health outcomes (diseases) and affects the ability to recover (susceptibility of mortality).

```
┌─────────────────────────────────────┐
│  MACRO or INSTITUTIONAL LEVEL        │
│  Social, institutional and economic  │
│  factors                             │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│     ENVIRONMENTAL INEQUALITY         │
└─────────────────────────────────────┘
```

COMMUNITY LEVEL
- *Built environment*

Traffic density
Housing quality
Water quality
Crowding
Neighbourhood quality (sanitation and waste)
- *Social environment*

Poverty concentration
Civic engagement
Crime and violence

INDIVIDUAL or MICRO LEVEL
- **Demographic factors**

Age, gender and race
Poverty
Educational status
- **Behavioural factor**

Health behaviour
- **Acquired factors**

Diet and nutritional status
Access to health care
Proximity to pollutants

Hazard → Emission → Exposure → Existing illness → Exposure-response function → Health effect → Ability to recover

**Community level indicators**

Individual level indicators

Figure 2:1: Multilevel model of the causes of environmental inequality
Sources: Kruize et al (2014), Morello-Frosch & Lopez (2006)and O'Niell (2003).

2.5.2. Community level

This level refers to the immediate context (the physical and social environment) that surrounds an individual (Soobader et al., 2006). The physical and social environment at this level produces a broad set of environmental risk. Risks tend to be distributed disproportionately across populations, communities and neighbourhood, accumulating in the neighbourhoods of low social groups (Kruize et al., 2014; Wright & Subramanian, 2007).

The physical environment produces risk factors such as: traffic burden, poor water quality, poor housing quality, poor neighbourhood conditions and crowding. These factors create hazards that when breathed into the body may disrupt the body's biological systems.

The social environment includes social risk factors such as: crime and violence, inactive civic engagement and poverty concentrated areas. Studies such as: Wright & Subramanian (2007) explain that these factors emit social pollutants that more or less act through the same mechanisms that potentially overlaps with those exhibited by physical pollutants and toxins. According to WHO Europe (2012), both physical and social risk factors cause psycho-social stress; a mental state that could adversely impair the immune and inflammatory systems. Thus, individuals living in poor social and environmental conditions have high exposure rates to the effects of psycho-social stress, making them vulnerable to negative health outcomes of environmental risk.

2.5.3. Macro/Institutional level

The macro level refers to a larger geo-spatial context that influence the direct conditions of an individual and the external environment that surrounds an individual (Soobader et al., 2006). They consist of those social, economic and political factors that create socioeconomic stratification and residential segregation (Kruize et al., 2014). According to Lopez and Frosch (2006), these factors shape the distribution of wealth and resources at the community and individual levels. This creates differing conditions at these subsequent levels, resulting in differences in exposure risk and health outcomes between social groups. These factors are subsequently discussed.

2.5.3.1. Economic factors

There are 3 basic economic factors that contribute to environmental inequality. These include: market economies, economic growth and urbanization.

Market economies

The capitalist economy is a strong driver of environmental inequality. According to Brulle & Pellow (2006), capitalism creates a treadmill of production; a process that generate a growing need for capital investment for the production of goods for sale in the market place. The expansion of the economy creates two elements: first, the creation of economic wealth, and second, the creation of the negative by-products of the production process. The impact is that social and economic benefits of the treadmill are unevenly distributed in favour of business and affluent communities, whereas the environmental risks and negative by-products associated with the treadmill are disproportionately concentrated among the

vulnerable poor groups with the least ability to resist the location of polluting facilities in their community.

Economic growth and urbanization

Two theories have shaped the understanding of the impact of economic growth and urbanization on inequality (Dinda, 2004; Kaika & Zervas, 2013). They are: The Kuznets curve (theory explaining the relation between economic growth and income inequality) and the environmental Kuznets curve (theory explaining the relation between economic growth and environmental degradation). Liu (2012) asserts that both theories result in an extreme injustice through the polarization of income and environmental quality. This process creates 3 population groups: the powerless and poor; the medium class and the powerful and wealthy. As the society urbanizes and economic growth increases, the poor and powerless groups may fall into deep economic and environmental poverty while the average and wealthy groups fall into moderate environmental poverty (Figure 2.2).
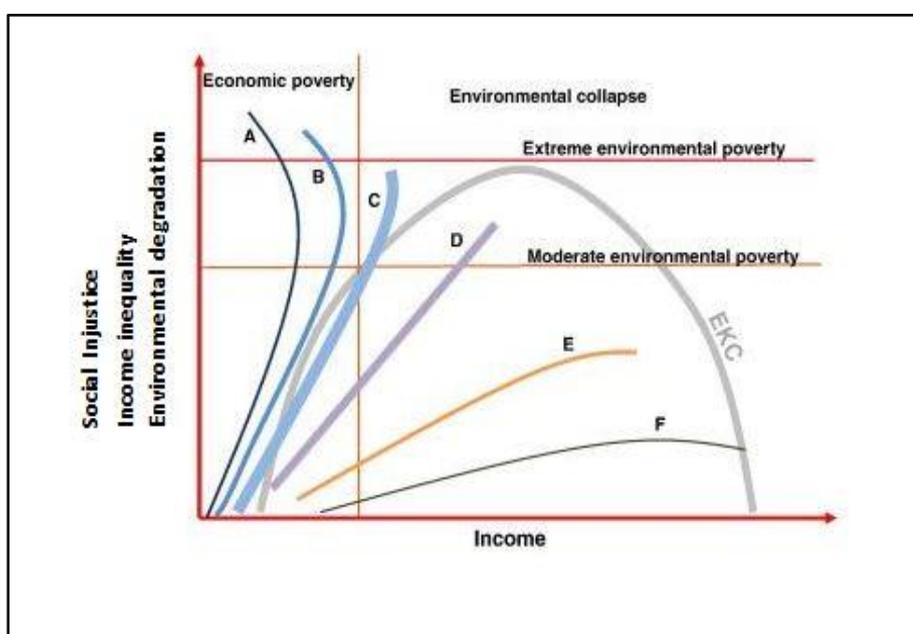
Figure 2:2: Level of inequality

2.5.3.2. Social factors

Factors such as migration and population growth have been identified as the social factors that create environmental inequality.

The influx of new migrants creates social differentiated residential communities(Szasz & Meuser, 2000). The tendency of migrants to cluster in a given location strengthens racial ties and creates racial segregated neighbourhoods.

Emigration also exerts the same level of influence on environmental inequality as immigration. Based on the theory of neighbourhood change and the push and pull model described by Kruize (2007), gentrification processes in urban centres might lead to the out

migration of wealthy households to sub-urban areas. Such movements, depreciates the value of land and attracts low income groups.

Population growth, often accompanied by economic growth could lead to shortage in housing supply. When population growth outpaces housing supply, a tight housing market is created. This implies that the cost of housing rises at a faster rate than income. This situation reinforces over time the tendency of a society to organize residential space along class lines. Affluent households find dwellings in high priced residential neighbourhoods and middle and low households in neighbourhoods with affordable prices (Szasz & Meuser, 2000).The resultant differences in the value of residential space could lead to the effects as described in the economic location theory, where the high cost of land in affluent communities discourages the siting of noxious facilities, attracting them to low cost, low income communities.

2.5.3.2. Institutional factors

The studies of Brulle & Pellow (2006) &Soobader et al. (2006) identifies institutional racism as a causal factor of environmental inequality. They explain that policy making institutions employ discriminatory mechanisms in environmental policy making processes. This influences landuse decisions wherein unwanted landuses are pushed away from wealthy neighbourhoods to poorer ones (O'Neill et al., 2003). The affluent social groups might also influence landuse decisions through their strong political will.

The Environmental inequality (EI) model of Pellow (2000) explains a multi-level influence in the formation of environmental inequality. Pellow explains that EI is not based on the unilateral influence of policy making institutions on poor social groups. EI emerges through processes of ongoing changes that involve negotiation and conflict amongst many stakeholders. Stakeholders that also include the poor and marginalized who make an outright choice to accept the unhealthy exposure they find themselves. Possibly because they gain from the risks they are exposed to by means of occupation and residence.

The theory of institutional racism is overruled by Szasz & Meuser (2000). They explain that environmental inequality is an outcome of unregulated development by planning authorities. Urban growth occurs at a rapid pace that is considered out of control for city planners to manage and coordinate. Over time, basic amenities depreciate in value, urban sprawl and slum persist and general degradation in the environment sets in.

2.6. Multi Differential Malaria Model

Based on the description made on the causal malaria risk factors and environmental inequality, we derive a model; the Multi-differential Malaria (MDM) model that explains the link between malaria and inequality. The MDM model selects malaria risk factors in Table 2.1. that are identical with the inequality factors mentioned in Figure 2.1. It further groups these factors into 3 levels:

- The macro level, which creates differences in the distribution and creation of malaria risk areas

21

- The community level, which creates differences in the exposure of malaria related hazard at a local level
- The individual level, which creates differences in the health outcomes of malaria in individuals.

Figure 2.3 provides the malaria risk factors for each of the 3 levels.

Differences in the distribution of malaria hazards
Social, economic and Institutional mechanisms

Differences in exposure (Community level stressors)

Differences in physical environmental conditions
- Housing quality and conditions
- Crowding
- Urban agriculture(vegetation)
- Water quality and availability (natural and polluted water surfaces)
- Neighbourhood conditions (waste dump, contaminated drainage channels)

Differences in social conditions
- Poverty concentration

Differences in health outcome (Differences in exposure and vulnerability) Individual level stressors

Differences in vulnerability
- Age(children, elderly, pregnant women)
- Poverty
- Low educational status
- Poor dietary and nutritional status
- Health status (existing illness and immunity)
- Health practices (use of clinical measures such as ITN and IRS)
- Distance to malaria risk sites
- Distance to health services

| Hazard | → | Plasmodium Virus | → | Exposure | → | Malaria | → | Mortality |

Effects of exposure differences

Effects of health outcome differences

Figure 2:3: Multi-differential malaria (MDM) model

3. STUDY AREA AND DATA

The discussion in this chapter splits into two main parts:

- Description of the study area with details on the administrative and topographic characteristics of the study area
- Description of data to be used for analysis where section 3.2.1. describes the ward level data and section 3.2.2. describes the individual level data.

3.1. Study area

The study is conducted in 2 Local Government Areas (LGAs) in Abia State, Nigeria: Aba South and Ugwunagbo. Both LGAs are located in Abia State situated in the south western part of Nigeria. Figure 3.1 shows the map of the study area in national, regional and local context. Ugwunagbo has an area coverage of 108km^2 and a population density of 904 inhabitants per km^2 (National Population Commission, 2010). It is made up of 15 administrative wards including: 2 urban wards, 7 sub-urban wards in the peripheral parts and 6 rural wards. Aba South on the other hand, has an area coverage of 49km^2 and a population density of 8650km^2. It consists of 14 urban administrative wards.
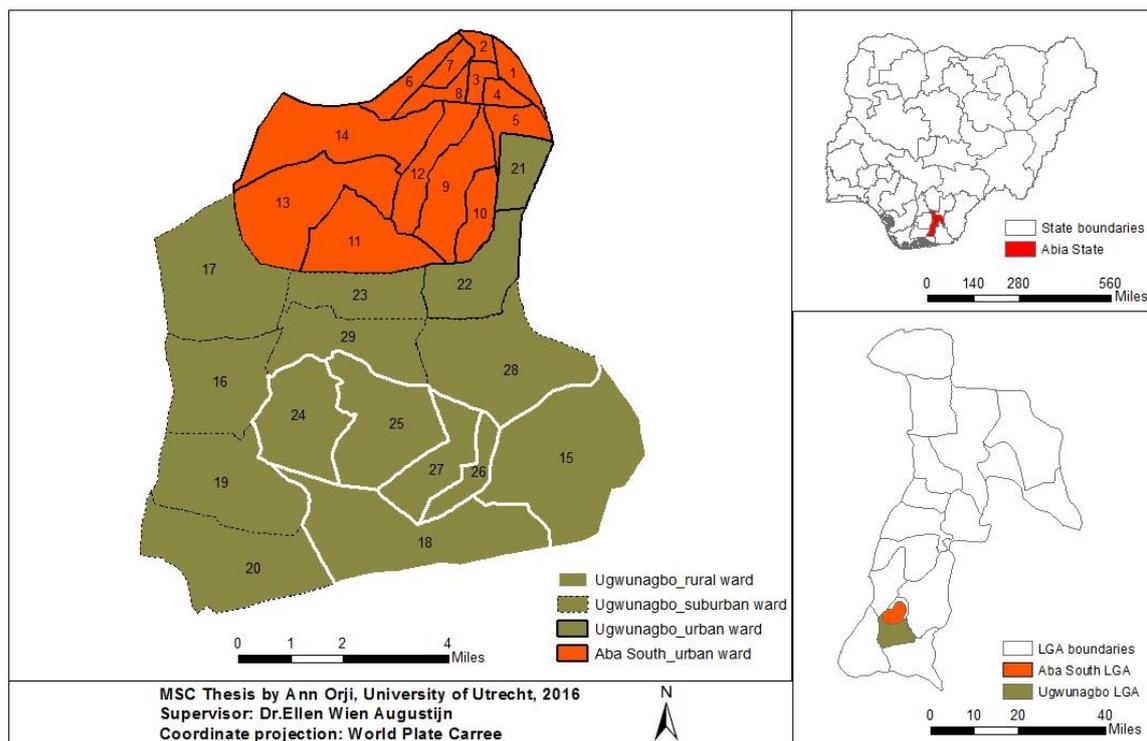

Figure 3:1: The study area in National, State and Local context

LGA	ID	Ward name	LGA	ID	Ward name
Aba south	1	Aba River	**Ugwunagbo**	15	Amapu Ideobia
	2	Aba Townhall		16	Ihie Obeaku
	3	Mosque		17	Umugo
	4	Enyimba		18	Obegu
	5	Igwebuike		19	Ihie Ukwu
	6	Eziukwu		20	Ngwaiyiekwe
	7	Asa		21	Nkpukpuevula
	8	Ekeoha		22	Owerri Aba
	9	Iheorji		23	Asa Umunka
	10	Umuogele		24	Umuarukwu
	11	Okporoenyi		26	Obeaja
	12	Ngwa		**26**	Umuchima
	13	Eluohazu		**27**	Umuada
	14	Glocester		**28**	Abayi Mbasaa
				29	Asa Amuhi

Table 3:1: Names of wards in Figure 3.1.

Economy

Ugwunagbo LGA is mainly dependent on agriculture; this classifies it as an agricultural economy. According to the Household Survey data collected for this study, about 70% of its inhabitants are farmers. The main agricultural products grown are food crops such as: yam, cassava, maize, cocoyam, rice and cash crops such as: oil-palm, cocoa, rubber and plantain. Aba South on the other hand, is the main commercial centre of the State. It is home to the largest market "Ariaria international market" in West Africa. The main occupation in this area is handcraft and trade.

Climate and topography

The climatic and topographical conditions of the study area are favourable for mosquito breeding. This makes the study area endemic for malaria transmission. Based on climate, Aba South and Ugwunagbo have an annual rainfall that ranges between 2000-2500 mm. According to Okezie et al. (2012) they receive more rainfall compared to the northern part of the state. The average temperature ranges between 22℃ and 31℃ per annum; a favourable temperature for vector breed and the extension of the life span of adult vector (Tanser et al., 2003). The study area is situated in a region with low altitude ranging between 19-65 m. The study area is surrounded by rivers and wooden green vegetation; habitats characterized as a suitable breeding site for mosquitoes. With a close distance to the river and a low altitude, the study area is vulnerable to the effect of flooding and its impact on malaria.

3.2. Data

This section is splits into two parts: the first part discusses the data used for the first level of analysis (ward level) and the preparation done on the data. It further presents an initial data exploratory analysis using the disease data contained in the dataset. The second part discusses the data collection processes for the second level of analysis (individual level).

3.2.1. Ward level data

3.2.1.1. Data content

Table 3.2 shows the data acquired for analysis and its details for the two LGAs. Analysis is conducted at two different levels: ward level and the individual level. The table contains data used for the ward level analysis; although, the data (river, health centres, breeding sites) are also used for the individual level analysis. The actual data used for the individual ward level analysis is obtained from the household survey; this is discussed in section 3.2.2.

Data	Source	Details for Aba South	Details for Ugwunagbo
Health data; malaria cases in (0-5) children 2010- 2015	Health unit of Ugwunagbo and Aba South L.G.A's	Malaria cases 2013-2015	Malaria cases 2010-2015
Population of (0-5) children	Demographic unit Ugwunagbo and Aba South L.G.A's	Aggregated to total population	Detailed showing total population of male and female
National, State and Local Government boundaries	Global administrative area http://www.gadm.org/country	Low resolution	Low resolution
Administrative ward boundary	Field work for collection of point features and digitizing with reference to natural features such as:roads	High resolution	High resolution
Buildings	Digitized using OpenStreetMap and random point created in ArcMap	Random points	Actual building points digitized fromOSM
River	Digitized from Topographic World Map	Detailed	Detailed
Mosquito breeding sites(puddles, ditches, sand excavation sites and waste dumpsite)	Field work	Present	Present
Health centre	Field work	Absent	Present

Table 3:2: Data, source and level of details for the LGA's

Both primary and secondary data were acquired. The primary data were sourced from field works and digitization made from the Topographic World Map and OpenStreetMap. These data include: administrative boundary, buildings, location of health centres and mosquito breeding sites such as: waste dumpsites and puddles. For the building data, the buildings in Aba South were not digitized because the area is densely populated with large number of buildings that are clustered together. We therefore used random points to represent the

location of buildings in this area; this is explained in details in Chapter 4. For the health centre data, this is unavailable for Aba South.

The secondary data were sourced from the health unit and the demographic unit of the two LGAs; Ugwunagbo and Aba South. The derived data are: the malaria data of 0-5 children from 2010 to 2015 and demographic data. These dataset are available for the year 2013 – 2015 for Aba South. Ugwunagbo has a complete dataset from 2010 – 2015.

3.2.1.2.   Data collection
The data collection for the ward level data was done through digitization and field work.

Map digitization

Due to non-availability of high resolution datasets, the following data were digitized using Aerial Imagery and ArcGIS 10.3.1:

Road and buildings

These features were digitized in OpenStreetMap (OSM); although this was done only for Ugwungabo LGA. Aba South was not digitized because it is a densely populated urban area which makes it time consuming to digitize. For Ugwunagbo, OSM was used because of its cost effectiveness and the possibility of uploading data in Shapefiles; a data format compatible with ArcGIS. The buildings digitized were limited to residential building, thus commercial and institutional buildings were not included. This exclusion was made because the activities of the population group being studied centre around their residential areas. The total number of buildings digitized sums to 12322. Figure 3.2.below shows a map of the digitized roads and buildings.
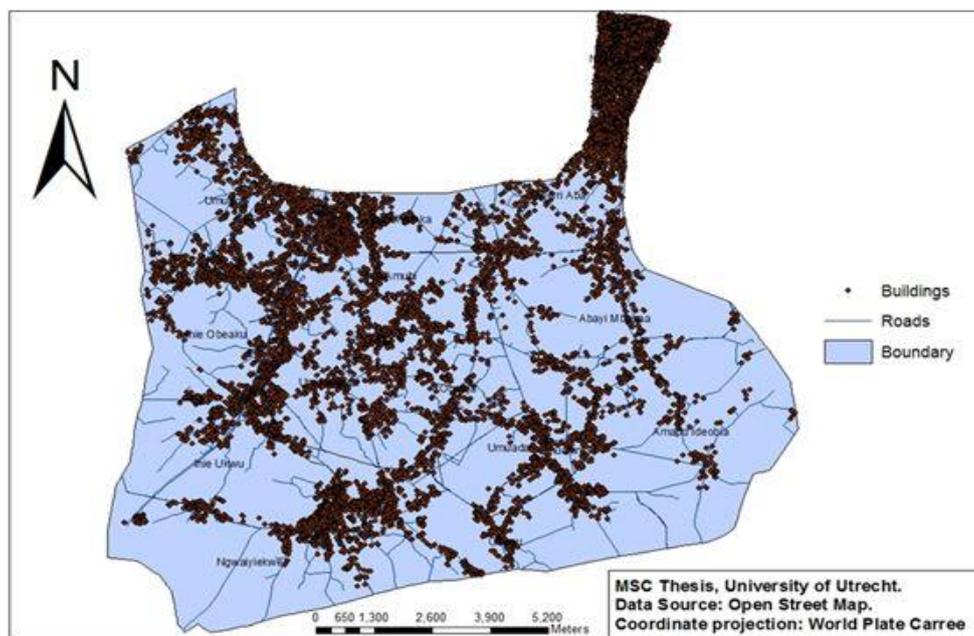

Figure 3:2: Digitized buildings and roads of Ugwunagbo LGA

Administrative ward boundary

The data collection process for this data was undertaken after the first field work was performed as described below. The Topographic World Map was overlaid with the road data and the point data of features collected during the fieldwork. Using the edit tool in ArcMap, boundary lines were traced around roads surrounding the point features. This was done for each ward until each ward boundary was derived. Since roads were not digitized for Aba South, the roads on the Topographic World Map were used as reference in the creation of ward boundaries for this area.

Field work

A four member team comprising of: two university students; a teacher and a local government worker were trained on the procedure undertaken during field work. The instructions for the fieldwork were laid out in a power point presentation and a record player. WhatsApp was used as a means of communication. The teacher was solely assigned the task of recording the data to ensure that the coordinates are collected in a standardized way. The local government worker piloted the entire field exercise by; driving the team around the study area and informing the local community about the purpose of the field exercise.

The field exercise was only done on the weekends; this to enable the team to carry out their regular work activity during the week days. The team went round each of the wards to spot the features of interest. For every ward visited, the team employed an indigene that helped in locating the exact spot of the feature being mapped.

The field work conducted by the team was done in 2 parts:

- *Field work for the collection of features for boundary mapping:* The aim of this field work was to get coordinates of identifiable features in the study area such as schools, health centres, villages and boundary point between wards. This was used to identify the boundary area for each ward which was then used as a reference data for creating the administrative ward boundaries. A GPS mobile device was used in the collection of points. The recruited team conducted this field work for Ugwunagbo LGA on the 19th and the 26th of June, 2016. Because, Aba South was included in the latter part of the study; features for this area were collected on the 25th of September 2016.
- *Field work for the collection of breeding sites and health centres:* The aim of this field work was to collect the coordinates of health centres and mosquito breeding sites which include: waste dumpsite and water surfaces - puddles, ditches and sand excavation sites. Fieldwork was undertaken in this area on the 10th and the 17th of July 2016. The collection of breeding site for Aba South was done on the same date; 25th of September when boundary features were collected. The location of health centres were only collected for Ugwunagbo as the detailed list of health facilities in Aba South was not provided. Figure 3.3 shows a map of the health centres in

Ugwunagbo and Figure 3.4 shows the map of mosquito breeding sites in the two LGA's.
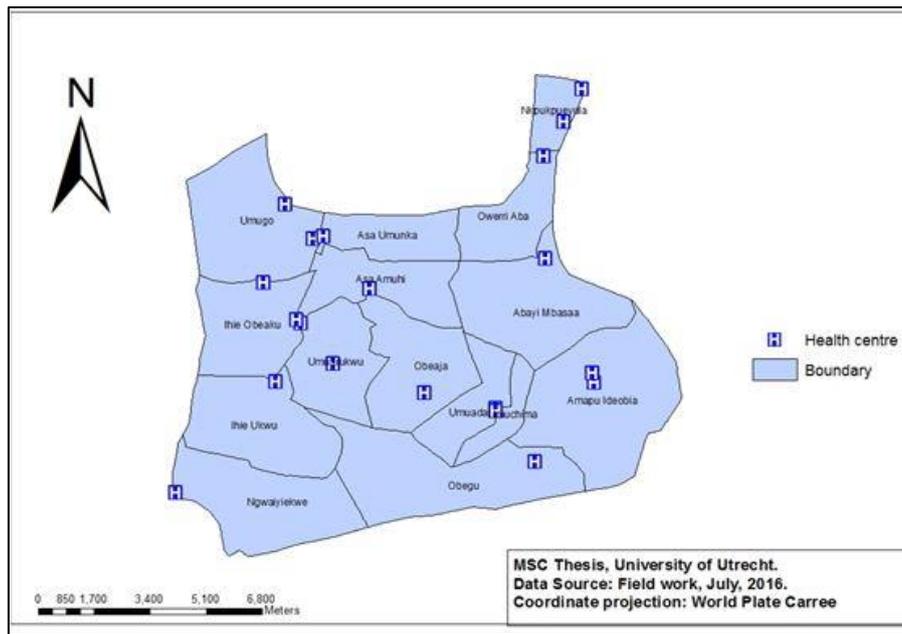
**Figure 3:3: Location of health centres in Ugwunabo LGA.**
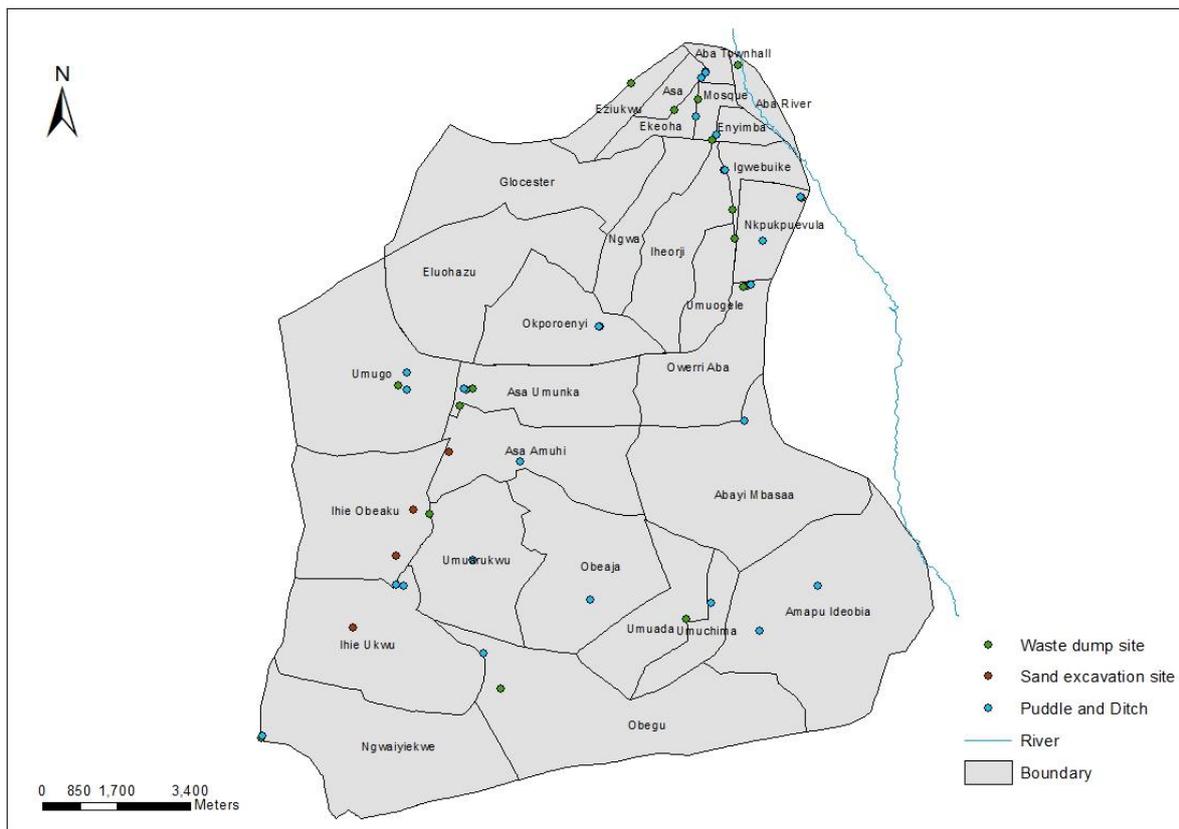
Figure 3:4: Malaria breeding sites in Aba South and Ugwunagbo LGAs

3.2.1.3. Initial exploratory ward level analysis

This section presents a pre-exploratory analysis of the ward level data (the malaria data for the years 2010 to 2015). It explains the temporal trend of childhood malaria as it occurred in both LGA' from 2010 to 2015. It further discusses the spatio-temporal trend of childhood malaria, which assesses the distribution of malaria occurrence by wards in the years being studied. It further assesses the distribution of malaria cases over the sexes. Because the data for Aba South lacks information on malaria occurrence from 2010-2012 as well as sex distribution of malaria, the complete data of Ugwunabo are used to explain sub-sections that require this information.

3.2.1.3.1. Temporal trend of child malaria

From the chart below (Figure 3.5), it can be observed that the number of child malaria cases in the study area has steadily increased from 2010 to 2015, although with exception of 2013 and 2015. For these exceptional years, Ugwunagbo had a decline in child malaria cases in 2013 while Aba South shows a decline in child malaria cases in 2015. The decline in malaria cases in Ugwunagbo is strongly linked to the health intervention programme that took place in the same year when treated bed nets were distributed to households. (Interview, HOD health; Ugwunagbo, 2015).
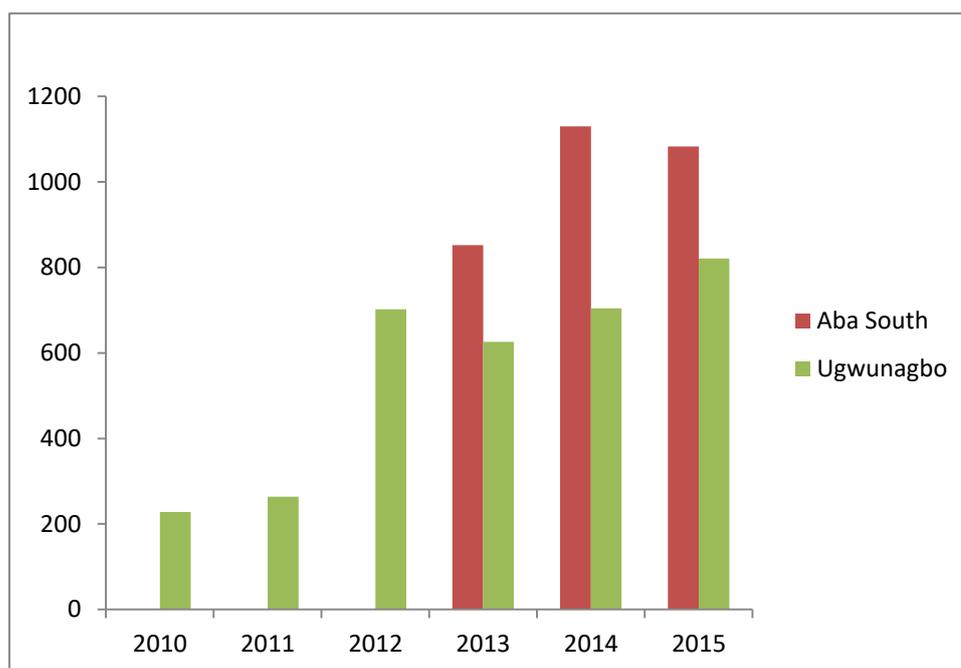
Figure 3:5: Total child malaria cases per 1000 population by LGA

The decline in malaria cases in Aba South in 2015 is in contrast to Ugwunagbo LGA where we see an increase in the number of child malaria cases. This difference cannot be explained by the influence of climatic factors such as rainfall and temperature; this because an influence of climatic factors should have resulted in a similar increase in both areas.

Figure 3.5 also points out a high level of malaria cases in both 2012 in Ugwunagbo LGA and 2014 in Aba South LGA. It can also be observed that the peak in malaria cases in 2014 in Aba

South had a similar effect in Ugwunagbo LGA where there was also an increase in malaria cases. This similarity could have been caused by the effect of climatic factors such as rainfall.

The reason for the peak in malaria cases in 2012 could be linked to the effects of flooding that took place in the country in the same year.
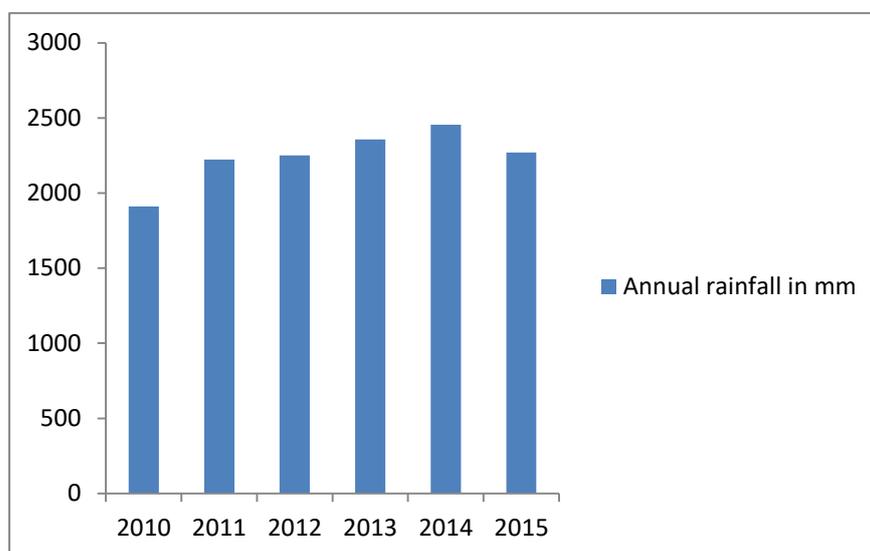

Figure 3:6: Annual rainfall of Aba in mm (NIMET, Alex (2013)).

It could be seen from Figure 3.6 that annual rainfall increased by 312 mm in 2011; followed by an increase of 29 mm in 2012. The increase in precipitation and the long rainy season experienced in 2012 (NIMET, 2013) led to flooding which gave rise to large water pools for mosquito breeding.

There was also an increase of 105 mm in 2013 and a further increase of 98mm in 2014. This might have accounted for the peak in the malaria epidemic in Aba South and Ugwunagbo in 2014 as earlier stated.

Another point to note in Figure 3.5 is the high number of malaria cases in urban areas (Aba South) compared to the sub-urban and rural areas (Ugwunagbo). This finding is different from those of Kigozi et al.(2015) and Robert et al.(2003) that validates rural areas to have higher malaria compared to suburban and rural areas. Our findings however, validates the notion of Austin (2014) and Ngom & Siegmund (2015) on the high level of malaria cases in urban areas where they explain causes to be income inequality and poor environmental conditions such as slums.

3.2.1.3.2.   Spatio-temporal trend of child malaria

The line graphs in Figure 3.7 below show the trend in the child malaria cases for the different wards in the LGA's. We can see two peaks as previously explained; the peak in 2012 in Ugwunagbo which affected almost all the wards and the peak in 2014 which also affected almost all the wards in Aba South and Ugwunagbo.

CHILD MALARIA PER 1000, ABA SOUTH WARDS

ADMINISTRATIVE WARDS OF UGWUNAGBO AND ABA SOUTH LGA

CHILD MALARIA PER 1000, UGWUNAGBO WARDS
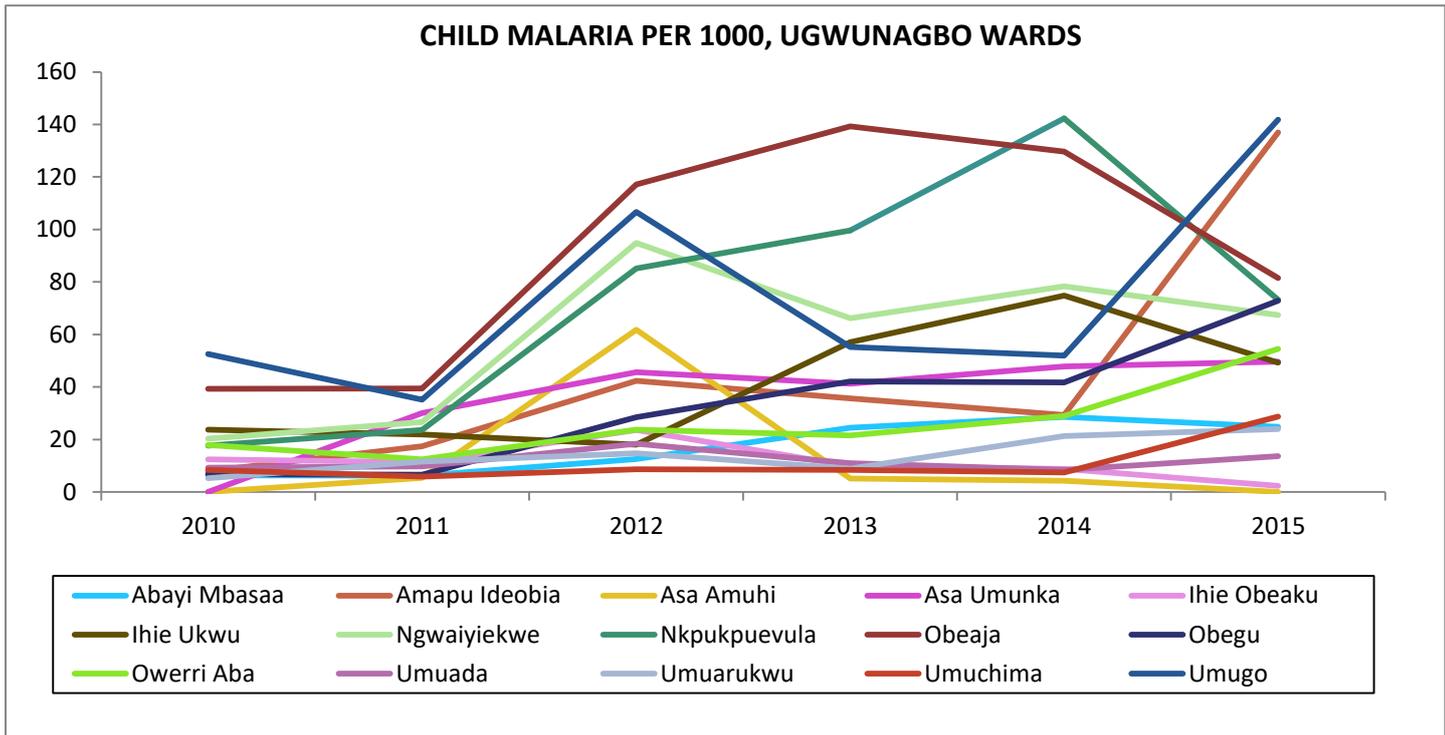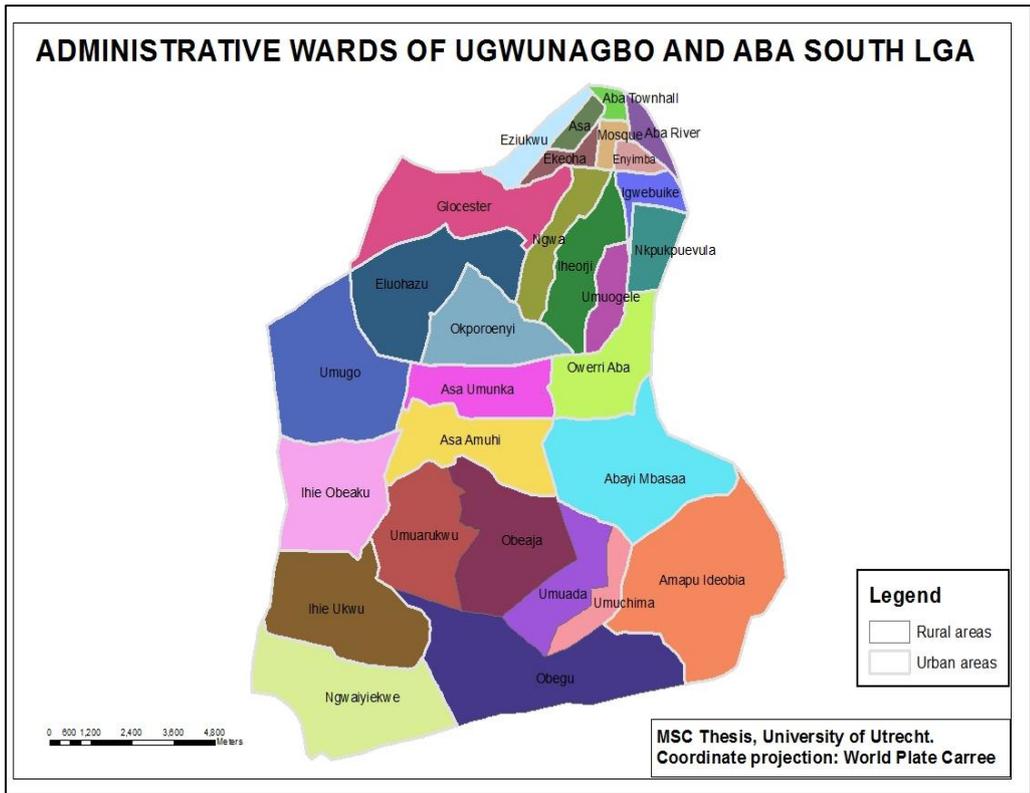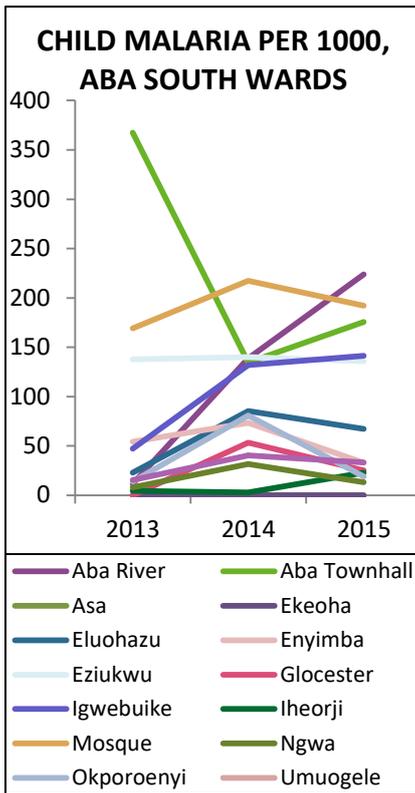
Figure 3:7: Child malaria cases per 1000 population of wards in the LGAs

It can be observed that the wards unaffected by the peak years such as "Ihie Ukwu" in Ugwunagbo and "Aba Townhall and Iheorji" in Aba South had an increase in malaria cases in the subsequent year; this was probably caused by the effect from the peak years.

The graphs draw attention on 5 administrative wards. The first two being; Nkpukpuevula and Obeaja in Ugwunagbo LGA, having a steady increase in child malaria occurrence from 2010 to 2014. The next being Aba River, a coastal area in Aba South which has had a steady increase in child malaria from 2013 to 2015. The last two wards being "Umugo and Amapu Ideiobia" in Ugwunagbo LGA which showed a marked increase in child malaria in 2015.
There are wards that had a relatively lower number of malaria cases compared to other wards. Wards such as: "Asa Amuhi, IhieObeaku, Umuchima, Umuarukwu" in Ugwunagbo LGA and "Iheorji, Ngwa and Glocester" in Aba South LGA.

It is therefore apparent from these findings that there is inequality in the prevalence of malaria per area. The burden of childhood malaria is uneven across the administrative units in the study area as certain areas bear a higher burden of malaria than other areas. Generally, children in the urban areas of "Aba Townhall, Mosque and Aba River" bear the highest burden of malaria as malaria rate range above 170 children per 1000 child population.

3.2.1.3.3. Sex distribution of child malaria

Figure 3.8 below shows the sex distribution of malaria in Ugwunagbo LGA. This chart is however based on the absolute number of malaria cases rather than relative number of malaria cases. This because the population data provided by the health department were aggregated and lacked information on population of male and female children. Aba South as pointed out earlier is not included in the analysis as it also lacks information on the number of malaria cases per gender population.
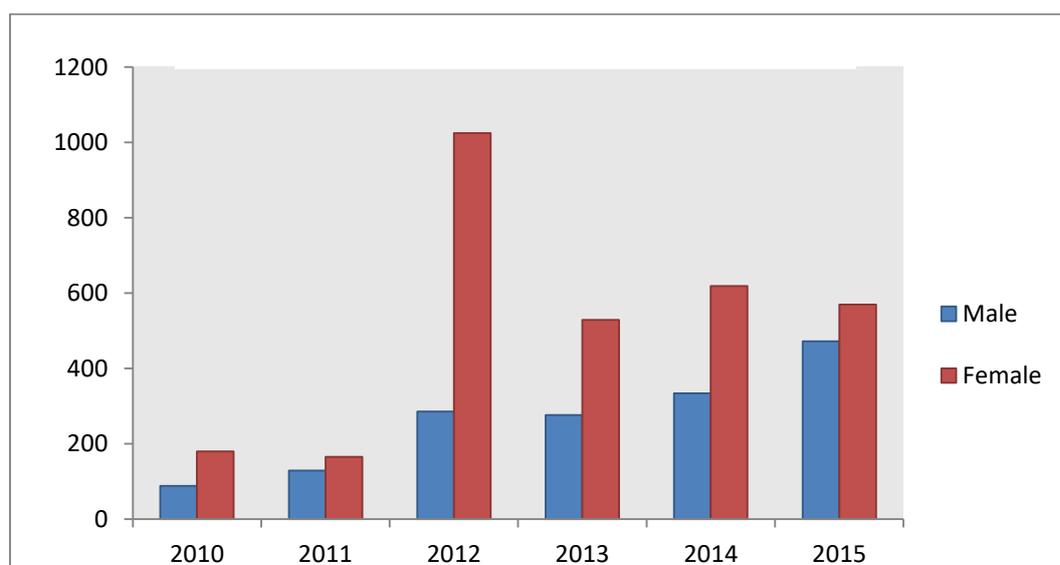

Figure 3:8: Sex distribution of child malaria in Ugwunagbo LGA

The chart shows that female children had higher cases of malaria than the males. This was the case in all the years from 2010 to 2015. We can also see a more than proportionate increase in female cases compared to males in 2012 which continued to 2014. This could have been as a result of the peak in malaria occurrence as has been previously explained and its effect in the subsequent year.

The vulnerability of female children to malaria is similar to the findings of Umeh (2013) conducted on Aba South, though his do not provide reasons for the higher prevalence of malaria cases in females compared to males. Based on the field observation done during the household survey, the field work team observed difference in the nature of outdoor play of female children compared to males. While female children are involved in outdoor play such as cooking with sand which requires them to sit in a place for a long period; the male children are involved in rolling tyre play which requires running around. Thus, the stationary and non-stationary positions might have a great impact on malaria transmission wherein female children staying in a stationary position during outdoor play are more likely to get mosquito bite than their male counterpart.

3.2.2. Individual level data

The individual level analysis uses data from the household survey conducted by the field work team. This section therefore describes the selection made for the survey and the methods used in the selection of the sample population.

Ward selection

The survey was limited to 3 administrative urban wards in Ugwunagbo LGA; Umugo, Asa Amuhi and Nkpukpuevula. These areas, based on the initial exploratory analysis in Figure 3.7 were selected because of the significant level of difference in malaria cases, compared to the other wards. Descriptively, Nkpukpuevula had a relatively high number of malaria cases for all the years, while Asa Amuhi had fewer malaria cases compared to other wards. The malaria cases in Umugo dropped significantly after the peak year in 2012.

Sample population

Prior to the survey, sample population and sample size was determined. This was to select a representative sample that was used for the survey. A questionnaire as can be seen in Appendix 1 was created, having all the required questions used for the individual level regression analysis. The variables derived from this survey are explained in Chapter 4.

The demographic data derived from the Population Unit of Ugwunagbo LGA were from 2010 to 2015. To derive the population for 2016 when the survey was done, a projection was made. The additive change projection method as described by Dennis et al. (2007) was used in estimating the 2016 population. This is represented as:

$$p_{0\,=\,P\triangle\,+pb} \qquad \text{equ1}$$

Where the projected population (p_0) equals the change in population (p$\triangle$) added to a base year population (pb).

Two base population figures from 2010 and 2011 were used. The population growth between the two years was raised to an exponent of 5; the number 5 being the number of years between the last base year, 2011 and the present year 2016. The figure derived was then added to the last base year, 2011 to derive the present population. See the table below for details.

Ward	Pop_2010	Pop_2011	Population growth(2010-2011)	Growth in 5 years	Estimate population (2016)	Sample size
Asa Amuhi	6332	6508	176	880	7388	95
Nkpukpuevula	9908	10183	275	1375	11558	96
Umugo	4843	4977	134	670	5647	95

Table 3:3: Population projection and sample size.

Sample size

The sample size calculator of the National Statistical Service of Australia was used to calculate the sample size of the estimated population. A 95% confidence level was selected with a 10% confidence interval. The result of this calculation is shown in the table above. This gives a total of 286 children to be sampled. This sum was rounded to 100 for each ward bringing the sample size to 300. The number of children sampled was selected randomly across the villages in each ward.

Sampling technique

A stratified random sampling was used for the household survey. This method was selected because the administrative wards have low aggregation units such as villages and clans. The villages making up the 3 administrative wards were classified as a group. The number of samples was evenly allocated across the villages and households were randomly selected from the village groups. The selection of households was influenced by network coverage, willingness of parents to participate and the availability of 0- 5 aged children.

The field team conducted the survey on the 2nd and 9th of August, 2016. Structured interview technique as described by Kumar (2014) was used to elicit information from the parents of children that were sampled. This involved a one on one interview with respondents, using the questionnaire in appendix 1 as a guide to structuring the questions being asked. This technique was chosen for reasons of: acquiring uniform information, getting an immediate response, reducing bias and explaining the questions clearly to respondents especially those from a low educational background.

Besides the use of interview and questionnaire, the study also used the observation technique. Section B of the questionnaire was completed using this technique. The outdoor play conditions between male and female children were also obtained using this technique.

4. METHODOLOGY

This chapter explains the analytical methods used in reaching the second research objective of:

- Identifying hot spot areas of child malaria incidence
- Assessing the variables in the MDM model (Figure 2.3) that are correlated with the identified hot spots and with malaria inequality in the entire study area

The analysis as can be seen in Figure 4.1 is performed in two aggregation levels: at a ward level and at an individual level. At the ward level, a cluster analysis is performed with the aim of identifying hot spot areas of child malaria occurrence. Subsequently, a regression analysis using Ordinary Least Square (OLS) model is done, which models selected ward level factors in the MDM model against child malaria incident rate. At the individual level, regression analysis is performed using the household survey data as described in section 3.2.2. The analysis at this level is constructed using 3 models: the OLS model, the Binary Logistic (BL) model and the Ordinal Logistic (OL) regression model. The outcome of these models is compared to derive the actual factors that explain inequality in malaria occurrence.
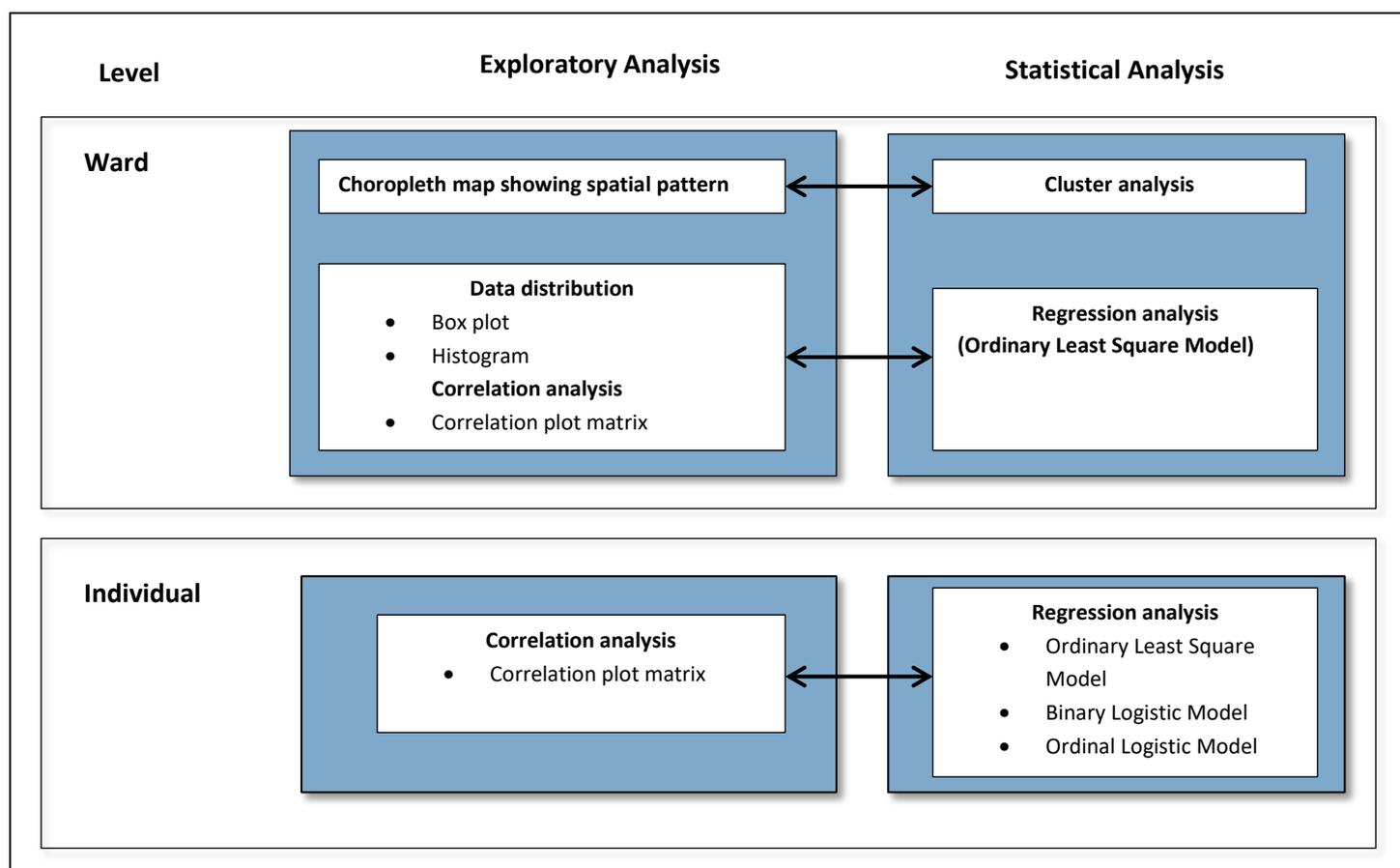

Figure 4:1: Analysis Framework.

For each of the analysis, an exploratory analysis is done on the data to assess the data distribution and to generate inferences that are validated in the actual analysis.

Although this research focuses on urban malaria, the variation between urban and rural malaria incident is explored at the ward level analysis. This because the study area consists of urban, sub-urban and rural areas and data is available for these areas.

The subsequent section discusses the variables selected for each level of analysis from the MDM model. It further discusses the 3 analyses: cluster, ward level regression and the individual regression.

4.1.   Variables and the MDM model

Table 4.1 below shows the variables that will be used for the regression analysis in relation to the factors in the MDM model.

In the **ward level analysis**, vegetation was not included as there seem to be a homogenous distribution of green vegetation across most part of the study area. Poverty concentration is not included in the analysis as there were no socioeconomic data available. Based on the data available, the variables: water, waste dump site, river, health centre and population density were used. Although population density is not included in the model, we assume it to have the same level of influence as the crowding risk factor (room occupancy rate) identified in the model. The distance is health centre is an individual risk factor but it is also measured at this aggregated level.

For the **individual level analysis**, two factors from the ward level inequality factors were included: housing quality/condition and crowding. For the housing quality and condition factor, housing quality factor was not included because there was similarity in the type of housing materials the sampled houses were built with. Materials being zinc roof and cement bricks which are resistant to mosquitoes (Lwetoijera et al., 2013). Housing condition was modelled as it varied across the sampled households. In this, the variables: roof, wall and window conditions were selected. For the crowding inequality factor, room occupancy rate; the density of persons per room were selected to be included in the analysis.

Level	Factors in MDM model	Variables selected for regression analysis
Ward level	Crowding	Population density
	Neighbourhood condition	• Waste dumpsite • Natural water surfaces (river) • Polluted water surfaces (ditches along the roadside, puddles and sand excavation sites).
	Distance to health service	Distance to health service
	Urban agriculture	

	Poverty concentration	
Individual level	Housing quality and condition	Housing condition • Wall condition (cracked or uncracked) • Roof condition(eave gap or no gaps) • Window (netted or unnetted)
	Crowding	Room occupancy rate
	Poverty	Income level
	Low educational status	Educational level
	Health practice	Use of preventive measure such as bed nets
	Distance to breeding sites	Distance to river, water and waste dump site
	Distance to health centre	Distance to health centre

Table 4:1: Selected variables and MDM model factor

The additional variables selected from the individual inequality factors include: income and educational level, use of preventive measures, distance to health centre and breeding sites.

A description on how these variables were calculated is detailed in section 4.2.1.2 and section 4.2.2.1.

4.2. Analysis

This section discusses the 3 analyses as shown in Figure 4.1. Two of the analyses: cluster analysis and ward level regression are performed at a ward level while the third analysis is the individual level regression done at an individual level. For each analysis, the methods used in computing the data, the exploratory analysis and the actual analysis are discussed.

4.2.1. Ward level analysis

The subsequent sections explain the two analyses done at the ward level; cluster analysis and the ward level regression analysis.

4.2.1.1. *Cluster analysis*

This analysis is done for the 29 administrative wards in Ugwunagbo and Aba South LGAs. It uses malaria incidence rate of 0-5 aged children to derive clustered hot spot of child malaria occurrence.

Rate(r) was computed as:

$$r = \frac{n}{p} * 1000 \qquad equ1$$

In which n is the total number of malaria cases per year and p represents the total population per ward.

For the fact that data is available for both LGAs in 2013 - 2015, analysis was limited to these years.

Prior to the actual analysis, an exploratory analysis was done using a choropleth map showing child malaria incidence rate (r) as input data. Choropleth maps has been found to be an effective visualisation tool to identify spatial clusters; this based on the grouping of similar colours across space (Zhang & Maciejewski, 2017). The classification method used is determined by plotting a graph of malaria incidence for the 3 years. The distribution of the data is then compared to the theoretical curve model to choose a suitable method.

Cluster analysis was performed using the Anselin local Moran's I statistical test in ArcMap. This test is chosen as it provides a more detailed analysis on spatial clustering compared to the global MoranI statistical test (Anselin, 2005). Four outputs are produced from this test: local moranI index; the Pvalue; Z score and the COtype.

$$Ii = \frac{x_i - \bar{x}}{\Sigma_i (x_i - \bar{x})^2} \sum_j w_{ij} (x_j - \bar{x}) \qquad equ2$$

The formula for the Local Moran's I index(Ii) is given in equation 2, where w_{ij} is the spatial weight between i and j; x_i and x_j are the values at location i and j and $\bar{x}$ is the mean of all the locations in the dataset. Because spatial weights are required for the calculation of the Local Moran's I index, the contiguity spatial weight was applied. This weight as described by Anselin (2005) is suitable for polygon dataset.

$$Z\ score = \frac{I_i - E(I_i)}{\sqrt{var(I_i)}} \qquad equ3$$

$$E(I_i) = -\frac{\sum_{j=1.j\neq i}^{n} W_{ij}}{n-1} \qquad equ\ 4$$

$$var(I_I) = E(I_I{}^2) - E(I_I)^2 \qquad equ\ 5$$

The P value and Z score measures the significance of the local moran's I index. Equation 3 to 5 gives the formula for Z score (Ping et al., 2004). The CO type comprises of two types: clusters and outliers. Hot spots are identified as a HH (high high) CO cluster type and are statistically represented by a positive local moran I index; a high Z score and P value less than 0.05.

4.2.1.2. *Ward level regression analysis*

This analysis uses the variables outlined in Table 4.1 to perform the regression analysis. We therefore used as dependent variable, the malaria incidence rate (r) of 2014 and 2015 as reported by the Health department and independent variables:

- Population density
- Mosquito breeding sites (waste dump site, water surfaces and river)
- Health centre

The subsequent section explains the method used in calculating the independent variables, the exploratory analysis and the regression model.

Calculation of variables

Population density was calculated with formula:

$$PD = \frac{\text{total population}}{\text{total area coverage in } m^2} (per\ ward)\ \text{equ6}$$

For the variables on breeding sites and the variable; distance to health centre, threshold distance was set for 200meters for waste dump site and water surfaces (Rosas-Aguirre et al., 2015). For river, a threshold distance of 1000m was set, this being the maximum flight distance of the Anopheles gambiae (Kaufmann & Briegel, 2004). This was used instead of 200 metres because none of the wards except Aba River (see Figure 3.4) are within 200 metres. For health centre, a threshold distance of 1000 metres was used.

The **total number of buildings** within the threshold distance of features was calculated for each ward. For Ugwunagbo, the digitized buildings were used. For Aba South, random points were created since building data was not digitized. This with reason that it is a densely populated area with clustered buildings; making it time consuming to have its buildings digitized. Figure 4.2 shows the model used to derive the points for Aba South. The value used as input was derived by:

$$Hld = \frac{\text{total population}}{5} (per\ ward)\quad \text{equ 7}$$

where Hld is the number of households and 5 is the average household size according to (Demographic Household Survey, 2013).

For precision, uninhabited areas were clipped off. Random points were then created on the inhabited areas.
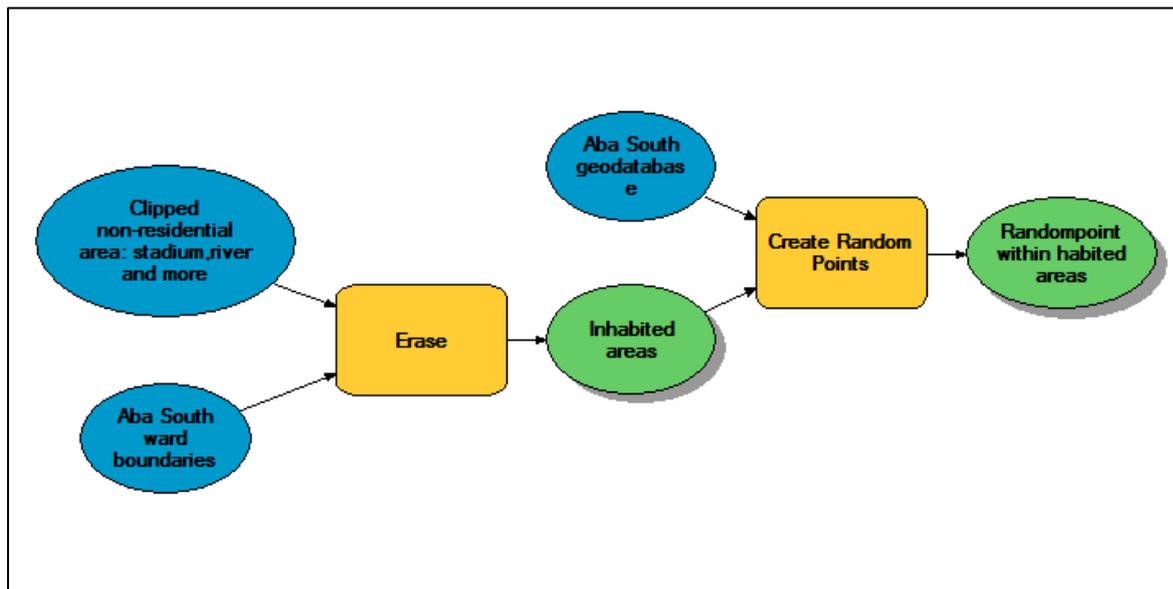
Figure 4:2: Model showing creation of random point for Aba South area
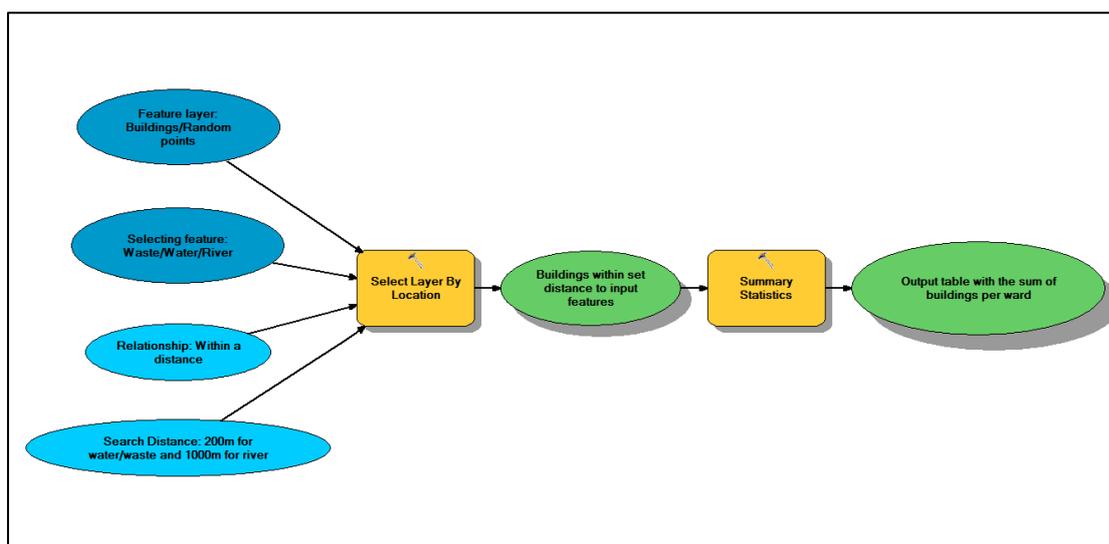
Figure 4:3: Model showing calculation of variables

The calculation of the variables is shown in Figure 4.3. It is expected that wards with high number of houses close to waste dump site; river and water surfaces will have high number of reported malaria cases. For the health centre variable; because we are using the reported cases of malaria from the Health unit, it is expected that wards with high number of households close to health centre will have high number of (reported) malaria cases.

Exploratory analysis

The exploratory analysis is done in two parts: on the data distribution and on the correlation between variables using R spatial statistical software.

The exploration on the data distribution is done using box plots (to identify extreme values or outliers) and histograms (to check for normal distribution). A histogram with a bell-like

shape indicates that the data is normally distributed while a histogram with a non-bell like shape indicates a skewed data distribution (Anselin, 2005). A skewed data distribution could either be positive or negative; positive when the tail on the right is longer than the left side and negative when the tail on the left is flatter than that on the right side. In the case of a skewed data distribution and outliers, a log transformation is done to normalize the data before regression is performed (Chan et al., 2014). This is done since a regression model assumes quantitative data to have a normal distribution.

For the exploratory analysis on correlation, a correlation plot matrix was used. This plot calculates the correlation coefficient of a pair of variables and represents them with colour. It has been found to be an effective tool in visualizing the correlation between multiple variables on a single plot (Mckenna et al., 2015). This plot is therefore useful in our analysis as there are multiple independent variables to be assessed against the two dependent variables as well as an assessment between the independent variables.

The correlation coefficient produced in the plot matrix ranges between -1 to +1. If the sign of the correlation coefficient is positive, then there is a positive correlation (Sedgwick, 2012). This is represented in blue colour and it for example indicates that wards with more houses within 200m of a waste dump site have a higher number of reported malaria cases.

If the sign of the correlation coefficient is negative, then there is a negative correlation. This is represented in a red colour and it indicates e.g. that wards with fewer houses within 200m of waste dump site and water surfaces are associated with higher numbers of malaria cases.

There is a perfect correlation with a correlation coefficient value of 1 or -1. A coefficient of zero indicates no association between or amongst variables.

Regression model

A univariate Ordinary Least Square (OLS) regression is fitted to test the significance of the correlation coefficient derived from the scatter plot matrix. The test is done using Pvalue which tests the null hypothesis of no correlation. Values below 0.05 reject the null hypothesis and therefore assume the correlation coefficient to be significant(Sedgwick, 2012).

Regression was fitted using malaria incident rate for 2014 and 2015 and independent variables:

X1: Number of houses within 200 metres of a waste dumpsite
X2: Number of houses per ward within 200 metres of a water surface
X3: Number of houses per ward within 1000m of a health centre
X4: Number of houses per ward within 1000 of the river
X5: Population density (number of people per m^2)

Equation 8 below shows the formula used in fitting the regression:

$$y_t = \beta_0 + \beta_1 X_n + \varepsilon_i \qquad \text{equ 8}$$

Where y_t is the dependent variable for a particular year; 2014 or 2015. β_0 is the intercept of the dependent variable; X_n is a given independent variable (X1, X2,X3, X4,X5); β_1 is the estimated value of the independent variables and ε_i is the error term. Since a univariate regression is applied, each of the equations was run separately for independent variables X1 to X5.

4.2.2. Individual level analysis

In this level of analysis, an individual level regression is performed using variables drawn from the household survey. Regression is performed using 3 models: OLS/Spatial model, binary logistic and the ordinal logistic regression model. Binary logistic and ordinal logistic regression are suitable model for categorical variables (Lwetoijera et al., 2013; Yusuf et al., 2010). While in the binary logistic model, the dependent variable is categorized into two levels (0 or 1), the ordinal logistic regression allows for more than 2 level. This makes these models appropriate for our analyses since all the data derived from the survey (except household size and room number) are of a categorical type. Although both models are suitable for our data and analysis, a comparison with OLS model is done to draw a good conclusion. The subsequent section describes how the variables were scaled for each of the regression models. It further describes the exploratory analysis and the three different regression models.

4.2.2.1. *Individual level regression analysis*
Categorisation of variables

Table 4.2 shows the variables and how they were weighted for the 3 regression models. A higher weight was given to a category with high chances of leading to malaria transmission and low weight to those with low risk of malaria transmission.

Variable	Category	Weight (Ordinary least square model)	Weight (logistic model)	Weight (Ordinal logistic model)
Wall condition	Uncracked wall	Cracked wall +open roof +unscreened window = 3 Closed roof + screened window + closed roof = 1 Combination of different housing condition=2	0	0
	Cracked wall		1	1
Roof condition	Closed roof		0	0
	Open roof		1	1
Window condition	Screened window		0	0
	Unscreened window		1	1
Malaria occurrence_ Independent variable	None	0	0	0
	Yearly	1	1	1
	Twice in year	2	1	2
	Quarterly	4	1	3
	Monthly	12	1	4
	More often	24	1	5
Education level	FSLC(Primary)	3		
	Waec/Technical(high school)	2		
	OND/NCE/BSC	1		
Occupancy rate	Household size(HS)	HS/HN		
	Household number(HN)			

Use of health measure	No measure used	5		
	Insecticide, coil	4		
	Pre-medication, traditional herb	3		
	Net	2		
	Net combined with other measures	1		
Income	20,000 and below	3		
	21,000 – 40,000	2		
	41- 60,000 and above	1		
Distance to water surfaces and waste dump sites	Numeric values	Actual distance of each sampled child to the nearest feature	≤200m=3 200 - 500m=2 ≥500m=1	
Distance to river	Numeric values		200 - 500m =2 ≥500m=1	
Distance to health centre	Numeric values		≤200m=1 200 – 500m=2 ≥500m=3	

Table 4:2: Variables and generated weights for each regression model.

For the dependent variable, weight for the OLS model was interpreted from the numeric values of the categories. In this, children in the yearly category (children who have malaria yearly) were weighted 1 and those who had malaria monthly where weighted 12 etc. In the binary logistic model, negative cases (children with no malaria) were weighted 0 while positive malaria cases where weighted 1. In the ordinal logistic model, malaria weight was assigned a scale of 0 – 5; 0 for those without malaria, 1 for those with yearly occurrence up to 5 for those who have malaria more often.

For income level; educational level; room occupancy rate and the use of preventive measures, the same weighted value was used for the 3 models. The reason for this is that the scaling done for the OLS model is already in a categorical form and could be used for the categorical models.

For the distance variables, the actual distance of each sampled child to the nearest feature was used for the OLS model. The calculation for the distances was done using the "near analysis" tool in ArcMap 10.4.1. For both binary and ordinal logistic regression model, distances were grouped into categories and weights were assigned for each category.

The threshold distance used in grouping the water, river and waste dumpsite variables was based on the grouping done by Ogunrinola & Adepegba (2012). Although adjustment was made on the first distance which was reset to 200metres according to the categorization of Rosas-Aguirre et al. (2015). Based on this, the distance to water and waste dumpsite was ordered into 3 groups; sampled children living <200metres were given a high weight of 3, this with reason that they are considered to be more vulnerable to contacting malaria. Those between 200-500metres were weighted 2 and those ≥500 metres were weighted 1. The distance to river variable was ordered into 2 groups because there were no sampled children within 200 metres of the river. Groups were set for: 200-500 metres with a weight of 2 and ≥500 metres weighted 1.

The distance to health centre variable was weighted differently as it is considered that children living close to health centres have lower cases of malaria due to access to effective health service such as: treated bed nets (Larson et al., 2012; Rutherford et al., 2010). Based on this, scale of 1 was given to children at close distance to health centre (200m); scale of 2 to those between 200 - 500m and scale of 3 to those living more than 500metres.

The grouping of these categories differs from the large threshold distances used by Feikin et al.,(2009) and Larson et al.,(2012). This because the maximum distance in the sampled data is 1315 metres; this with variation between the sampled wards (Nkpukpuevula (636 metres), Asa Amuhi (674 metres) and Umugo (1315 metres)).

These short distances are linked to the fact that health centres are at close proximity to the sampled wards especially for Nkpukpuevula. It can be seen from Figure 4.4 that the maximum distance of the sampled wards (wards with blue dots) to the nearest health centre is 1000 metres for Nkpukpuevula and 3000 metres for Asa Amuhi and Umugo.
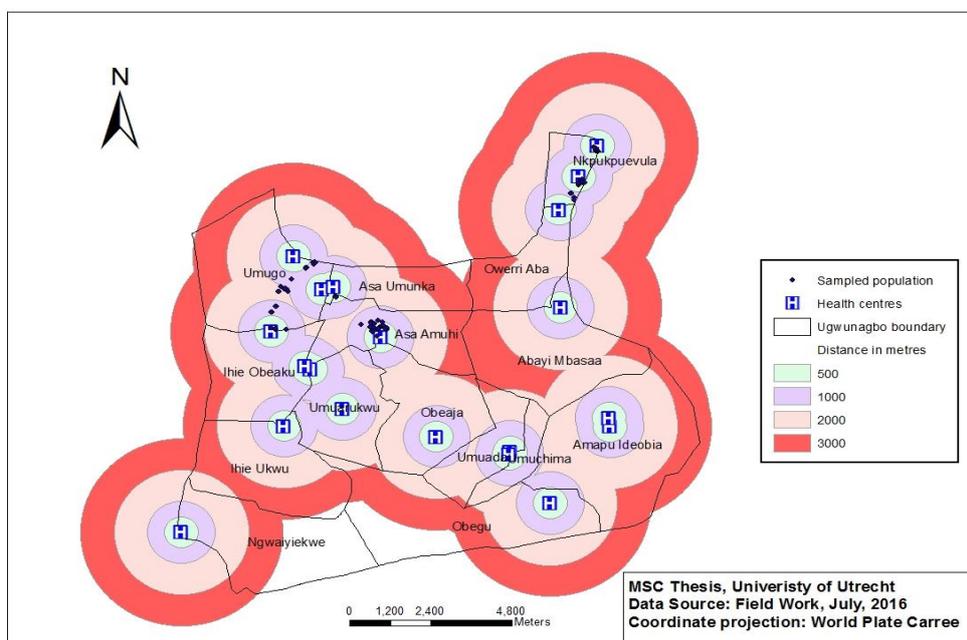
Figure 4:4: The distance of Wards from the nearest health centre

Also, due to poor network coverage experienced during field work, the areas in Asa Amuhi and Umugo within 2 to 3000 metres were not sampled. Although for this ward, households at the North west part were sampled but as can be seen from Figure 4.4 these households are close to the health centre of an adjacent ward, Asa Umunka.

The findings of Feikin et al. (2009) show a difference in the impact of health centres on the population living within 500m and 1000m. Thus, despite the short distances applied in our analysis, we still expect a difference in the impact of health centres on the sampled population living more than 500metres from the health centre.

Exploratory analysis

This analysis uses a correlation plot matrix to assess the independent variables that are associated with child malaria frequency. Because the data used in this analysis is of a categorical type, an assessment on outliers and normal distribution is not required.

Correlation plot matrix as described in section 4.2.1.2. shows the correlation coefficient; a measure used to determine the strength of association between pairs of variables. Coefficient with positive value shows a positive correlation which for our analysis indicates, for example: that sampled children from low income are associated with a higher number of malaria cases. This association should be understood in the context of the weighting done on the variables as shown in Table 4.2. High risk categories such as low income were weighted with high values while those with low risk were given low values. So, explaining a positive correlation (a high to high effect), we interpret a relationship between a **high** weighted category (which literally is a **lower** category such as: low income) and a high malaria occurrence. Based on this, we interpret a positive correlation between malaria and income to be children from low income background (weighted with high value) is associated with a high number of malaria cases.

A negative coefficient value indicates a negative correlation. This implies that: sampled children from high income and educational background (having low weights) are associated with a higher number of child malaria cases.

Coefficient value of 1 or -1 indicates a perfect correlation, while the value of zero indicates no correlation.

Regression model

Regression using the household survey data was performed to test the significance of the correlated coefficient derived from the coefficient plot matrix. It uses the frequency of malaria occurrence per sampled child as the dependent variable and the following as independent variable:

P1 Wall condition
P2: Roof condition
P3: Window condition
P4 : Room occupancy rate
P5: Income level
P6: Educational level
P7: Use of preventive health measures
P8: Distance to water surfaces
P9: Distance to waste dump site
P10: Distance to river
P11: Distance to health centre

A univariate regression was fitted for the 3 models. In this, regression was repeated for the dependent variable and each independent variable. Regression was done in 4 parts: for all the 300 sampled children and for each of the 100 sampled children in Nkpukpuevula, Umugo and Asa Amuhi. The formula for the regression models is described below.

The formula for the Ordinary Least square (OLS) model is similar to equation 8 in section 4.2.1.2., but using the variables outlined above. To derive an appropriate spatial model to correct for the effect of spatial autocorrelation in the OLS result, a multivariate regression was run in "Geoda"spatial statistical software. In the derive result, an assessment was done on the spatial dependency test. A significant spatial lag (Pvalue<0.05) suggests running a Spatial lag model while a significant spatial error suggests running a Spatial error model (Anselin, 2005). The suitable spatial model is run in Geoda.

The **binary logistic (BL) model** is fitted with formula:

$$\text{Logit (p)} = \beta_0 + \beta_n P_n + \varepsilon_i \qquad \text{equ 9}$$

$$\text{Odds} = \frac{p}{1-p} \qquad \text{equ10}$$

Where p is the predicted probability of a child being malaria positive. The odds of the same child having malaria is given in equation 10. β_0 is the intercept; P_n is the independent variable (P1 - P11) ; β_n is the regression coefficient of each independent variables and ε_i is the error term (Yé et al., 2006).The probability of having malaria from the equation is either 1 or 0. This therefore implies that the dependent variable should have a binary data type. This is the reason why the grouping done on the dependent variable (for the BL model) as can be seen in table 4.2 was done in a 0-1 scale.

The **ordinal logitisic(OL)** regression model is fitted with formula:

$$Y_j = \beta P_n \qquad \text{equ11}$$

Where Y_j is the coefficient of each category (0 - 5 in Table 4.2) of the dependent variable with ranges between 1 to -1. βP_n is the coefficient of each explanatory variables (Ye, 2013).

For either models, a positive coefficient and a significant p-value (<0.05) indicates that an increase in the explanatory variable is strongly associated with an increase in malaria occurrence. A negative coefficient and a significant pvalue indicate that an increase in the explanatory variable is significantly associated with a decrease in malaria occurrence.

5. ANALYSIS AND RESULT

The first research objective has been described in the previous chapter by:

- Describing both vector and host based factors that influence the occurrence of urban malaria
- Explaining in detail the concept of environmental inequality and causal factors of inequality using the multi-level model
- Deriving a Multi-differential malaria (MDM) model; adjusted from the multi-level model which links the factors of environmental inequality with that of urban malaria.

This chapter answers the second research objective. It uses the data explained in Chapter 3 and the analytical methods described in Chapter 4 to assess hot spot areas of child malaria occurrence. It further analyses at a ward and individual level, the environmental inequality factor from the MDM model that are correlated with the malaria hot spots and with child malaria occurrence in the entire study area. This chapter therefore provide results of:

- Cluster analysis using the Anselin spatial autocorrelation in ArcMap which identifies malaria hot spots in the study area from the year 2013 – 2015.
- Ward level and Individual level regression analyses.

For each of the analysis, an exploratory analysis is performed to provide insight on the actual analysis.

5.1. Cluster analysis

This section discusses the cluster analysis performed using Anselin Spatial Autocorrelation in ArcMap. This analysis identifies hotspot areas of child malaria occurrences for the years 2013-2015. Before the cluster analysis is discussed, an exploratory analysis is done using choropleth map which maps the malaria incidence rates for the 3 year period. This is aimed to identify potential clustered areas and to validate the result of the actual cluster analysis.

5.1.1. Exploratory analysis; spatial distribution of child malaria

A graphical presentation as shown in Appendix 3 was done to select a suitable classification method. The graph shows a geometric distribution of child malaria cases per 1000 population. This suggests the use of a geometric interval classification method according to the theoretical classification curve.

Figure 5.1 below shows the result of the spatial pattern of child malaria cases classified into 5 classes using the geometric interval classification method. It can be observed from the map, a clustering of child malaria in the riverine areas in the north-eastern part of the study area; a pattern that recurred throughout the years. Conclusions might be drawn; explaining distance to river to be a cause for the clustering of child malaria in the northeast. This might not be the case as there are other riverine wards in the northeast such as: Asa and Ekeoha and Amapu Ideobia in the southeast that have dissimilar pattern (low malaria cases). This further suggests a factor or a number of factors other than nearness to the river that explains the clustering pattern of malaria. This is tested in the regression analysis in 5.2.

Besides, the clustering observed in the northeast, another clustered pattern can be seen in 2013 in the southwestern part of the study area. Although this pattern did not persist in the subsequent year; it however recurred in 2015. This pattern also suggests other factors than nearness to river to be the cause of clustering of child malaria.

To statistically test areas that are clustered hot spot of malaria occurrence, a clustered analysis is performed.

Figure 5:1: Spatial pattern of malaria, 2013-2015.

5.1.2. Result of cluster analysis

Figure 5.2 below shows the result of the cluster analysis performed using Anselin local Moran's I test in ArcMap. This analysis identified 4 malaria hotspots in the north-eastern part as can be seen in Table 5.1 with a high z score and a significant pvalue<0.05.

Cluster type	Wards	2013		2014		2015	
		Zscore	pvalue	zscore	pvalue	zscore	pvalue
Hotspots	Aba River	2.1341	0.03283	3.028	0.0025	3.3531	0.0008
	Aba Townhall	2.7916	0.0053	2.6596	0.0078	3.3763	0.0008
	Igwebuike			2.2208	0.0264		
	Mosque			3.2973	0.0001	2.6749	0.0075
Outliers High-Low	Obeaja			-2.3042	0.0212		
Outlier Low High	Asa	-2.0872	0.0369	-2.9432	0.0033	-2.5657	0.0103
	Ekeoha			-2.1656	0.0304		

Table 5:1: The z score and pvalue of clustered areas.

No cold spots were identified in the analysis; however outliers were detected, indicating the presence of areas with low malaria occurrence. Three outliers were identified as can be seen in Table 5.1. The first two are Asa and Ekeoha; wards with lower malaria cases, surrounded by malaria hotspot areas. The third outlier is Obeaja, in the rural southern part of the study area. This ward is an area having high malaria cases. It is surrounded by wards with low malaria epidemic; Umuchima and Umuarukwu as can be seen in Figure 5.1.

Figure 5:2: Cluster map of child malaria incidence

Generally, there are 2 hotspots that recurred throughout the 3 year period; Aba Townhall and Mosque.

The result from this analysis gives a strong conclusion on the presence of inequality in malaria. It can be deduced that children living in the north eastern part of the study area are more susceptible to malaria than children in other wards.

5.2. Regression analysis

This section provides the regression results at the ward level (5.2.1) and individual level (5.2.2). For the ward level analysis, the descriptive statistical analysis is performed using box plot and histogram. Correlation plot matrix is used for both ward and individual level analysis.

5.2.1. Ward level regression analysis

At this level of analysis, regression is performed using variables interpreted from the community level inequality factors in the MDM model. Table 4.1(methodology chapter) lists the variables considered for this analysis. Thus, for the **dependent** variable, the child malaria incidence rates for 2014 and 2015 is used; rate being the total number of child malaria incidence per 1000 0-5 population and for the **independent** variables the following are analysed:

- Number of houses within 200metres of a waste dumpsite
- Number of houses within 200metres of an artificial water surface
- Number of houses within 1000 metres of the River
- Number of houses within 1000metres of a health centre
- Population density

A strong correlation between these variables and malaria indicates that wards with high number of households close to these features have high malaria cases than those wards further away.

5.2.1.1. Descriptive statistical analysis

Prior, to the regression analysis, descriptive statistics is done on the data using box plot, histograms and correlation plot. Box plot is used to identify outliers in the data, histogram is used to check for normal distribution and correlation plot is used to check for correlation amongst the variables.

Data distribution

Figure 5.3 below shows the data distribution for each variable across the upper and lower quartile range. The figure on the left shows the data distribution for each of the explanatory variables and that on the right shows the data distribution for malaria incidence rate in 2014 and in 2015.

Table 5.2 shows the statistical values used to compute the box plot. This includes for each variable: minimum value, maximum value, median (Q2), mean and Q1 (lower quartile) and Q3 (upper quartile).
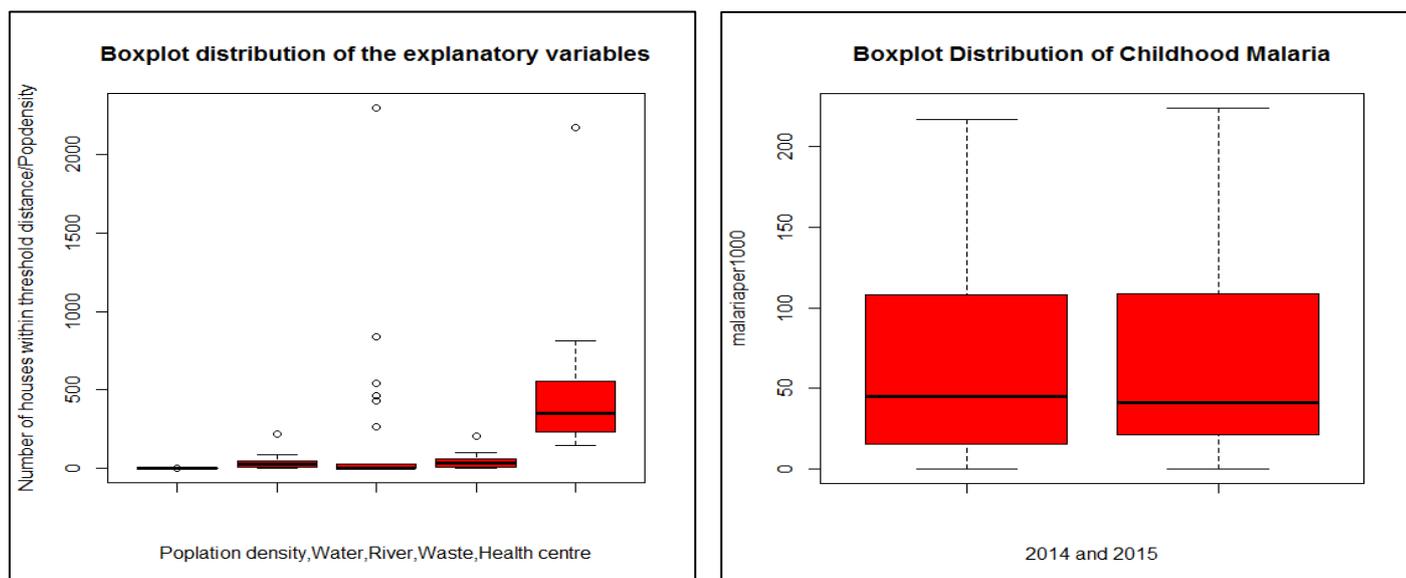
Figure 5:3: Box plot of independent (left) and dependent variables(right).

Variables	Min	Max	Median	Mean	Q1	Q3
Malaria rate 2014	0.00	217.30	45.00	62.72	18.00	54.40
Malaria rate 2015	0.00	224.00	41.20	65.65	21.95	95.45
Distance to water	0.00	219.00	25.00	33.25	2.75	41.50
Distance to river	0.00	2300.00	0.00	174.50	0.00	22.25
Distance to waste	0.00	201.00	34.00	41.57	7.75	56.00
Distance to health centre	145.0	2177.0	348.5	504.1	230.8	541.5
Population density	0.0000900	0.2960000	0.0021850	0.0252868	0.0003225	0.0088625

Table 5:2: Statistical values for box plot computation.

It can be observed from the box plot on the left of Figure 5.2, that there are 4 variables; water, health centre, waste and river with outliers (points above the last line on the plot). The dependent variables; malaria incident rate 2014 and 2015 on the right of Figure 5.3 has a no outliers.

Variable	Outlier(wards)
Distance to water	Nkpukpuevula
Distance to river	Aba River Aba Townhall Enyimba Igewbuike Mosque Nkpukpuevula
Distance to waste	Nkpukpuevula
Distance to health centre	Nkpukpuevula

Table 5:3: Outliers and Wards

It can be seen in Table 5.3 that the ward "Nkpukpuevula" recurred as an outlier for all the 4 variables. This is explained by the highest number of buildings/households that are close to breeding sites (water, river and waste) compared to other wards. The fact that the same ward is an outlier for the health distance variable suggests effective health coverage for the population in Nkpukpuevula. This could be explained by the fact that there are a higher number of buildings close to health centre in the ward compared to others.

The presence of outliers gives an indication of a skewed data distribution. Histograms were plotted (Figure 5.4) to check for skewedness. It can be seen from the histograms, that the 4 variables with outliers have a skewed distribution.
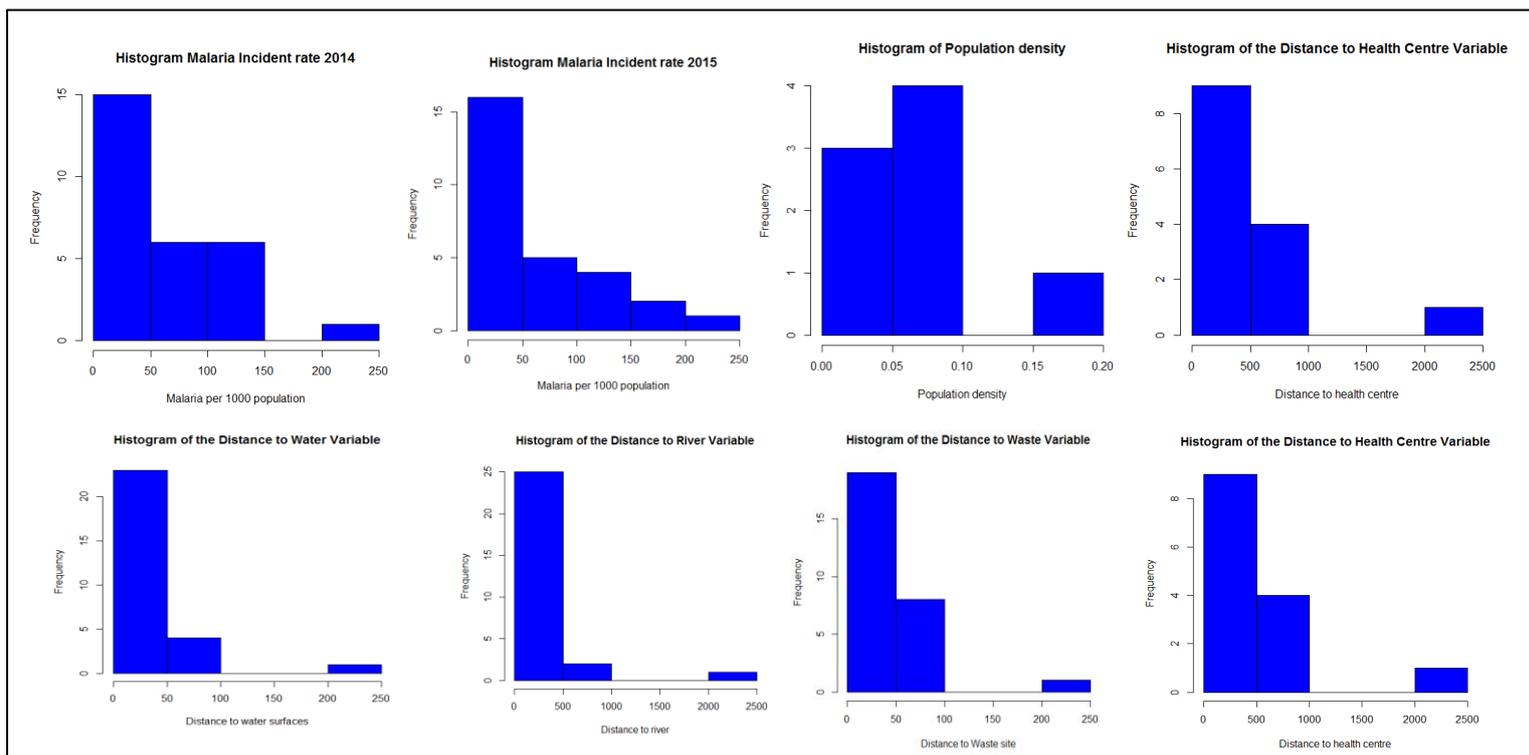
Figure 5:4: Histogram showing data distribution of variables

Although, the dependent variables seem to have a normal distribution from the box plot analysis in Figure 5.3; the histogram shows these variables have a positively skewed distribution. Nevertheless, the population density variable has a normal distribution shown in both box plot and histograms.

Because regression analysis assumes data to have a normal distribution, we normalized the variables with skewed distribution; water, waste, river, health centre, malaria incidence rate for 2014 and 2015, using a log transformation with formula:

$$Log10 \ (v+1);$$

where v is the data value in each variable. The value of 1 was added to each variable to avoid the zero values in the data from getting infinity values. The log transformed data were used for both correlation analysis using correlation plot and the actual regression analysis. Both analyses are discussed in the subsequent section.

Correlation analysis

Figure 5.5 shows 4 different correlation matrices of the dependent variables and the explanatory variables. The first plot shows matrix of all the wards, the second plot on the top right shows correlation matrix for Aba South, the third plot on the bottom left shows correlation for Ugwunagbo urban area and the 4th shows plot for Ugwunagbo rural area. In the matrix for the combined wards and Aba South, the distance to health centre variable was not included because health centre data is not available for Aba South. However, this
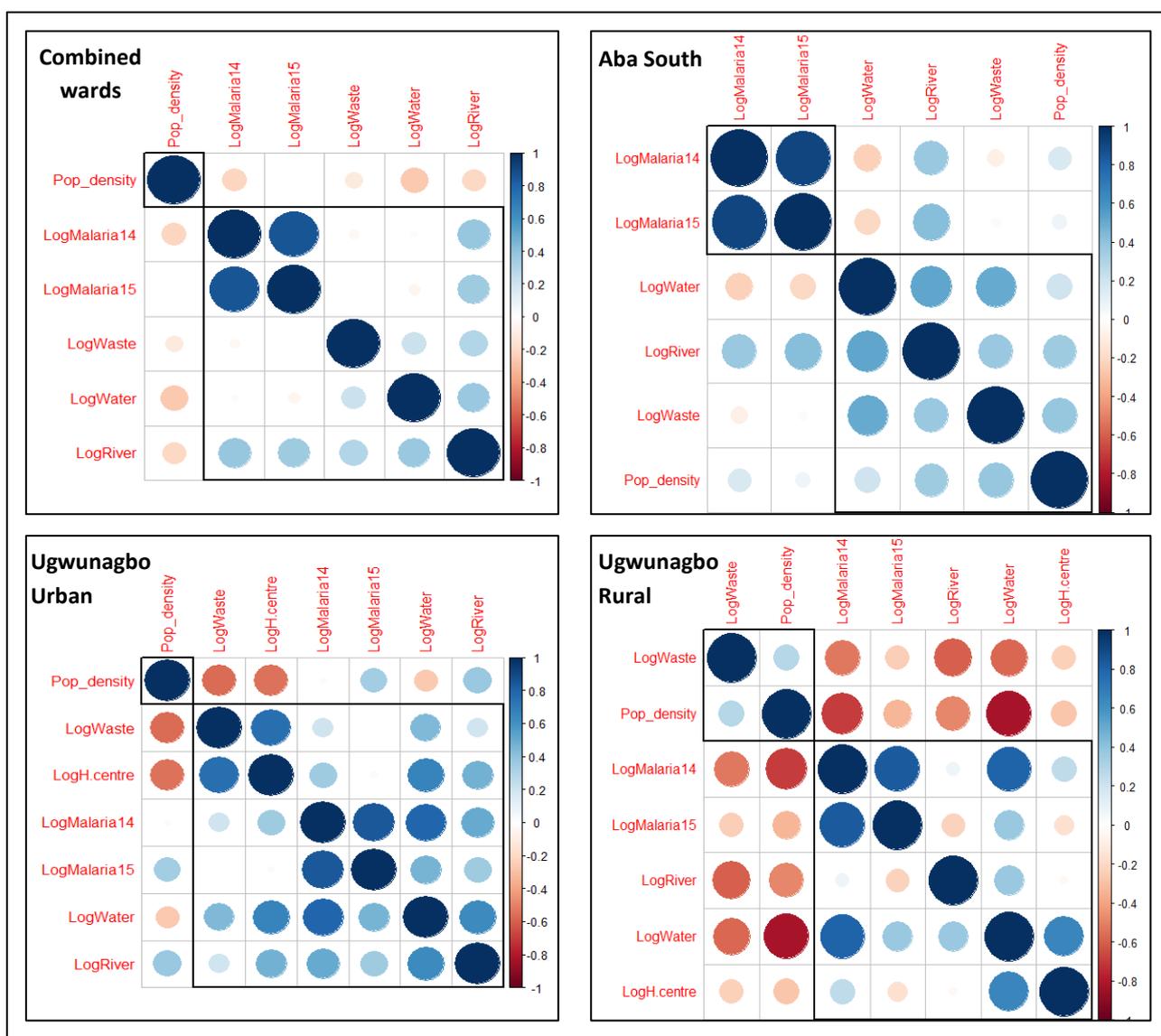
Figure 5:5: Correlation plot matrix of the aggregated wards and sub areas.

variable is plotted in the matrix of Ugwunagbo Urban and rural area where data is available. Interpreting the plots, the bar on the right of each plot, scaled from negative 1 to positive 1 represent the correlation coefficient. The size of the circle indicates the strength of the correlation with 1 or -1 indicating a strong correlation and values close to zero indicating no correlation. The blue coloured circle indicates that the correlation is positive: larger values in the independent value correlates to a larger value in malaria. The red coloured circle

indicates that the correlation is negative: smaller values in the independent variable correlate with high values in malaria.

The strength of correlation is assessed by the size of the circles. The bigger the blue or red circle, the stronger the correlation and the smaller the blue or red circle, the less correlation there is. However, the blue circle with the most intense correlation with value 1 is the correlation between a variable matrixed against itself and should be ignored.

For all the plots, it can be seen that there is a strong correlation between malaria incidence in 2014 and in 2015. This suggests that wards that had high malaria cases in 2014 were also the same wards with high cases in 2015. There are differences in the results of the plots; differences between the aggregated wards and the individual sub areas and differences in result between the years. The individual plot for the sub areas provide more details on correlated variables than the combined ward plot. The subsequent paragraphs give description of each plot matrix.

In the **combined ward** matrix, it can be seen that river shows a positive weak correlation with malaria incidence in both 2014 and 2015. This indicates that wards with higher number of households living within 1000 meter of the river are more likely to have malaria than those living far away. All other explanatory variables as can be seen in the plot have no correlation with malaria.

In the matrix of **Aba South**, it can be seen that river and water are correlated with malaria. The plot shows that river has a positive influence on malaria for both years (when more people living within 1000 metres of the river, the number of malaria increases). Water has a negative correlation with malaria, meaning that when less people living within close distance to the water, malaria incidence increases. This is unexpected.

Besides the correlation between malaria and river/water, the plot shows positive correlation amongst the environmental variables and population density. This suggests that the wards close to river are the same wards close water and waste. This is true, as it can be seen in appendix 6 which shows a clustering of these environmental variables in the North east of Aba South situated close to the river.

In the matrix of **Ugwunagbo urban area**, 2 variables; river and water have a strong positive correlation with malaria. Water is shown to have a very strong correlation compared to the other variables. Waste on the other hand has a very weak correlation compared to the other variables. In 2015, water, river and population density show a positive correlation with malaria. However, the strength of correlation with respect to river and water is not as significant as it was in 2014.

In the correlation matrix of **Ugwunagbo rural area**, we see a strong positive correlation with water and a strong negative correlation with population density and waste dumpsite. Water is shown as a correlated variable for both years, although the strength of correlation is more in 2014 than in 2015. This is similar to the findings in Ugwunagbo urban as previously described

The negative correlation between malaria and population density indicates that more malaria occurs in wards with a lower population density. This is not unexpected as malaria is traditionally assumed to occur in rural areas(De Silva & Marshall, 2012). It could be that less dense rural areas are surrounded by vegetation which are breeding sites for mosquitoes; making people living in these areas more vulnerable to malaria.

The analysis indicates that malaria show different correlations in urban versus rural areas. In urban areas, malaria is most correlated to river and water. In rural areas malaria occur in the least populated areas. The difference between the patterns in urban versus rural areas explain why the plot for the combined wards (having both urban and rural areas) shows the least correlation. Combining rural and urban areas seems to average out the patterns.

It can also be concluded that the two years 2014 and 2015 showed very similar correlations. This indicates a stable pattern.

5.2.1.2.   Regression results

Table 5.4 shows the Ordinary Least Square (OLS) regression result of child malaria incident rate 2014 and 2015 and explanatory variables. The regression analysis was done in 4 parts: the first part performed regression with all the 28 wards in Aba South and Ugwunagbo; the second part with the 14 wards in Aba South urban area; the third part with the 8 urban wards in Ugwunagbo area and the final part with the 6 rural wards in Ugwunagbo area.

Malaria prevalence in these areas was tested against the independent variables distance to river, population density, distance to water and distance to waste. All analyses were conducted for 2014 (left side of table 5.4) and for 2015 (right side of table 5.4). Pvalues test the null hypothesis of no correlation. Values below 0.05 reject the null hypothesis and therefore assume the variables to be correlated. In table 5.4, significant values are indicated with an asterisk.

It can be seen from the table, that correlation was found between malaria and the number of people living close to river for the combined ward analysis in both 2014 (0.0285) and 2015 (0.0364). However, for the different areas, river is only significant for Aba South and for the year 2015 (0.0382). There is also correlation between malaria incidence and water for the year 2014 (0.0478).

Area	Variable	Estimate	Std.Error	Pvalue	Estimate	Std.Error	pvalue
		2014			2015		
All wards	River	0.259182	0.108691	0.0285*	0.27095	0.12197	0.0364 *
	Population density	-2.1812	1.95949	0.3258	0.22085	2.14693	0.9190
	Water	-0.290587	0.238505	0.2910	-0.19997	0.20279	0.3343
	Waste	-0.143656	0.162824	0.4322	-0.09064	0.18270	0.6245
Aba South Urban	River	0.3820	0.1830	0.06645	0.4115	0.1696	0.0382 *
	Population density	18.2797	69.8605	0.79947	-23.0192	64.7376	0.7303
	Water	-0.5291	0.3215	0.13420	-0.5821	0.2979	0.0825
	Waste	-0.1121	0.3941	0.78250	0.1210	0.3652	0.7479
Ugwunagbo Urban	River	-0.6515	2.6794	0.1558	0.0000	0.0000	0.447
	Population density	0.0000	0.0000	0.1017	0.0000	0.0000	0.268

	Water	0.0000	0.0000	0.0478 *	0.0000	0.0000	0.177
	Waste	0.0000	0.0000	0.4083	0.0000	0.0000	0.526
	Health centre	0.0000	0.0000	0.8459	0.0000	0.0000	0.959
Ugwunagbo Rural	River	0.0573427	0.37915	0.88711	0.398601	0.242729	0.17590
	Population density	-3.20817	1.5353	0.10488	-2.22928	1.45645	0.20061
	Water	0.724677	0.278981	0.06020	0.465397	0.29753	0.19282
	Waste	-0.21	0.391436	0.62007	-2.22928	1.45645	0.20061
	Health centre	0.086	0.54203	0.88162	0.318	0.420029	0.49114

Table 5:4: OLS regression result of Wardlevel analysis

The result is similar to the findings in the correlation matrices where river was found to be significant for the combined ward and Aba South analyses; also, for Ugwunagbo where water was found to be significant in 2014. Dissimilarity lies in the correlation between malaria and water in 2014 for Ugwunagbo Urban. In the regression result these variables are insignificant while the correlation matrix shows them to be significant. This is explained by the weak correlation in 2014 as shown in the correlation matrix compared to 2015 and also, the close significant pvalue of 0.06645 in the regression result. The same applies to Ugwunagbo rural area where the correlation matrix shows water to be correlated with malaria in 2014 and the regression result show a close to significant pvalue of 0.06020.

Based on this result, conclusions can be drawn on 3 key points:

- The significant influence river has on child malaria occurrence in the hot spot areas identified in Figure 5.2. This is seen in the high pvalues (0.0382) in 2015 and close to significant value (0.067) in 2014 for Aba South.
- The different environmental factors correlated to child malaria in the administrative areas. This is seen by the strong correlation, river has with child malaria in Aba South compared to Ugwunagbo urban areas where water is strongly significant.
- The differences in the correlation of environmental causal factors overtime where in Aba South in 2015, river had a significant influence on child malaria and in 2014 was not significant. This is also the case for Ugwunagbo where in 2014 water was significant to child malaria, but was not significant in 2015. The insignificance of water in 2015 could be related with the decrease in precipitation in the same year as shown in which must have led to a decrease in the volume of water in puddles and ditches.

5.2.2. Individual level regression analysis

This section uses the household survey data to model the factors that are significant to child malaria. Factors as it relates to the individual level factors of the MDM model. We therefore, use as dependent variable, the **frequency of child malaria occurrence** and independent variables:

- Income level
- Educational level

- Use of preventive health measures
- Room occupancy rate
- Housing condition (wall, roof and window condition)
- Distance to breeding sites (water, river and waste dump site)
- Distance to health centre

Regression is fitted with three models: OLS, Binary logistic and Ordinal logistic regression. Details on how the variables were ordered for each model are provided in Table 4.2(methodology section). Although our discussion is based on the Ordinal logistic regression, a comparison is done with the other models to draw variables that are significant to child malaria.

5.2.2.1.   Descriptive Statistics

Prior to the regression analysis, descriptive statistics is done on the data using scatter plot. There is no need to check for outliers and skewedness since the data is of an ordinal data type; ordered during data preparation. Figure 5.6 shows a correlation plot matrix of the dependent variable and the explanatory variables. This uses the data categorized for the ordinal logistic regression model as shown in Table 4.2.

As previously explained, the circles in red show that variables have a negative correlation while the circles in blue, shows a positive correlation between the variables. The strength of the correlation is assessed by the size of the blue or red circles. The value of 1 or -1 shows a very strong correlation. However, the coefficient with the value 1 is the correlation between a variable matrixed against itself and should be ignored.
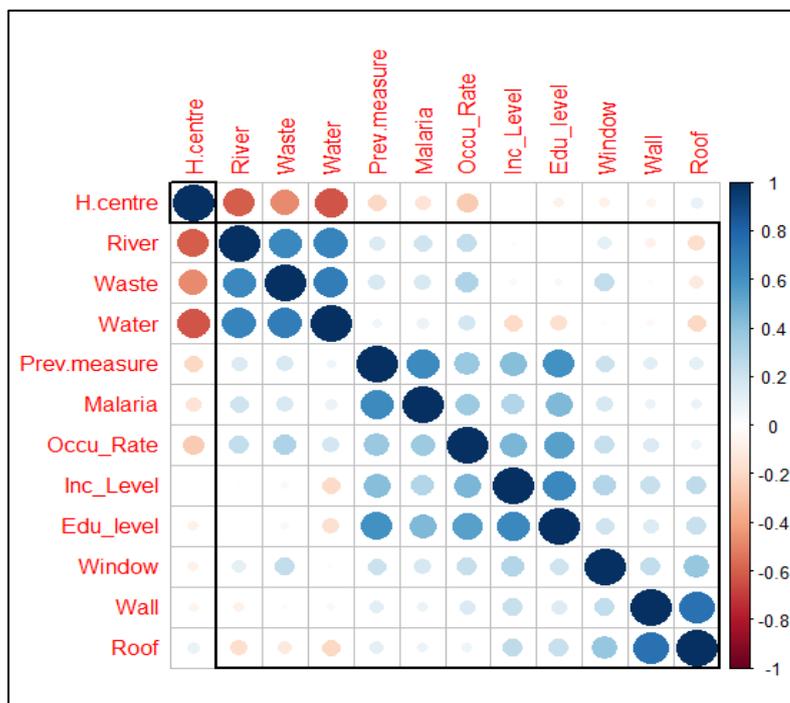

Figure 5:6: Scatter plot matrix of Malaria frequency and explanatory variables

To interpret the plot, it is important to understand how the variables were categorized and weighted as detailed in Table 4.2(methodology chapter). Categories that lead to high risk of malaria transmission were given high weights. For example; children from low income where given high weight of 3 and those from high income homes were given low weight of 1.

Assessing the correlated variables with malaria, we can see a positive correlation between malaria and the variables: use of preventive health measure; education level; income level and room occupancy rate. Explaining this relationship, we assert that children who do not use any preventive health measures (scaled highest value 5) were more likely to have malaria than children who used nets or a combination of net and other measures (this with a scale of 2 and 1 respectively). With educational level, children with caregivers from a low educational background (scaled with the highest value 3) were more vulnerable to malaria compared to those with caregivers having higher educational background (scaled with low values of 1). The same is explained for income level, children from low income homes (scaled with high value of 3) had more chances of having malaria than those from average and high income background. In addition, the correlation between malaria and room occupancy rate implies that children in households with high number of persons per room had more chances of having malaria than those from households with less number of occupants in a room.

For the environmental variables, it can be seen that river, waste and water have a positive correlation with child malaria, although the correlation is not as strong as the variables mentioned earlier. There seems to be a higher level of influence of waste and river on malaria than water. This relation implies that children living close to mosquito breeding sites (especially river and waste) are likely to have more malaria cases than those living far away. This is expected as several findings such as: De Silva & Marshall (2012) and Rahman (2006) proofs this to be true.

There is a weak negative correlation between child malaria occurrence and distance to the health centre. This implies that children living close to the health centres are likely to have more malaria than those living far away. This is unexpected as the findings of Feikin et al. (2009) and Rutherford et al. (2010) proof that children living close to health centre have better access to health service and are less likely to have malaria than those living far away. This could be as a result of the dependency on other health services and measures outside government funded health centres. Services such as: drug/pharmaceutical shops and traditional herbs as identified during the survey.

For housing conditions, we can observe a relatively stronger positive correlation between malaria and window condition compared to wall and roof conditions. There is also a positive correlation between window condition and the use of preventive measures. It could be explained that children from households who use effective health measures such as bed nets may not have netted windows.

Besides the correlation between malaria and the explanatory variables explained above, there are correlations amongst the explanatory variables. We can observe a correlation amongst the environmental variables; river, waste and water which is similar to the findings drawn from the ward level correlation matrix analysis. The positive correlation between "occupancy rate and these environmental variables" strongly suggests the clustering of crowded households to malaria breeding sites. Furthermore, since there is a positive correlation between both educational and income level and occupancy rate, it could be explained that children with low income/educational caregivers reside close to mosquito breeding sites than others. This makes them more vulnerable to malaria than children from a wealthy background.

The positive correlation between "income level and housing condition" strongly suggests that children from low income homes are more likely to reside in homes with poor housing condition than those from average and high income homes.

Conclusively, we can deduce from the correlation plot matrix that the variables; use of preventive measures, room occupancy rate, wall condition, income and educational level and distance to river and waste are correlated factors to child malaria occurrence. However, assessing on the strength of the correlation we can conclude that the use of preventive measures, occupancy rate and income and educational level are the key significant factor to malaria. This is validated by the regression analysis in the subsequent section.

5.2.2.2. Regression results

Table 5.5, Appendix 5 and 6 shows the regression results for this level of analysis. Table 5.5 shows the regression result of the Ordinal logistic model. Since this is the most suitable regression model for our data, our discussion is based on this model. However, comparison is done with the result of the other models; the Ordinary Least Square (OLS) Regression and the Spatial model in Appendix 5 and the Binary Logistic (BL) Regression Model in Appendix 6.

Spatial error model was modelled as a result of the spatial dependency test performed on the (OLS) result as shown in appendix 4. This test suggests the need for spatial correction using the spatial error model (SEM) since spatial lag is insignificant while the spatial error is significant (Anselin, 2005). We therefore fitted a spatial error model to correct for the effect of spatial autocorrelation in the OLS regression result. Both OLS and the corrected SEM result are contained in the table in Appendix 5.

For the regression results, analysis was done in 4 parts: for all the wards and for the individual wards; Nkukpuevula, Asa Amuhi and Umugo. The frequency of malaria occurrence for each of the 4 groups was tested against the explanatory variables. The results for each analysis show both estimate coefficient value and the pvalue. The estimate coefficient value signifies the rate of change; positive or negative of the dependent variable; malaria frequency when the explanatory variable increases. Pvalue test the null hypothesis

of no correlation. Values below 0.05 reject the null hypothesis and therefore assume the variables to be correlated. In the tables, significant values are indicated with an asterisk. More asterisks signify a high level of significance.

Variable	All wards		Nkpkpuevula		Asa Amuhi		Umugo	
	Estimate	Pvalue	Estimate	pvalue	Estimate	pvalue	Estimate	pvalue
Distance to the nearest health centre	-0.3863	0.0059***	-0.2756	0.3055	-0.5820	0.3344	-0.5820	0.3344
Distance to the river	1.1566	0.0004***	-0.2506	0.5562	More than 1000 m from river with similar values and cannot be modelled			
Distance to the nearest water surface	0.2365	0.1326	0.0699	0.7953	0.5820	0.6029	0.5820	0.3344
Distance to the nearest waste dump site	0.4121	0.002***	0.0699	0.7953	-0.1048	0.6476	-0.1048	0.6476
Occupancy rate	0.5374	0.0001***	0.3087	0.0222*	0.5166	0.0072**	0.5166	0.0072***
Wall condition	0.4191	0.1627	-0.3925	0.5537	1.6322	0.0013***	1.6322	0.0013***
Roof condition	-0.1345	0.766	-0.4725	0.5557	1.6237	0.0019***	1.6237	0.0019***
Window condition	0.6857	0.0013***	0.0637	0.8829	0.3536	0.3392	0.3536	0.3392
Income level	0.9601	<0.0001***	1.1927	0.0132**	0.5243	0.0342*	0.5243	0.0342*
Educational level	1.4393	<0.0001***	1.1987	0.0001***	1.1477	0.0018***	1.1477	0.0018***
Use of health measures	1.0037	<0.0001***	0.8567	<0.0001**	0.9894	<0.0001***	0.9894	<0.0001***

*Level of significance

Table 5:5: Ordinal logistic Regression result of Individual level analysis

From the Ordinal logistic result in Table 5.5, it can be seen from the aggregated ward result that all the explanatory variables except distance to water, wall and roof condition are significant to child malaria occurrence. However, the variables; distance to river, waste, health centre and window condition are not significant in the different individual ward analysis. The same difference can be seen in wall and roof conditions were for the two individual wards; Asa Amuhi and Umugo these variables are significant but are not significant in the aggregated ward analysis. The subsequent section discusses in detail the results of each explanatory variable and possible reason for the difference in the aggregated ward and the individual ward analyses.

Comparing models

Similar results were obtained from the Ordinal logistic (OL) regression in table 5.5 and that from OLS/Spatial error model in Appendix 5. The only difference is in the individual wards

where the OLS model has occupancy rate to be insignificant for Umugo and distance to water significant for Asa Amuhi.

The result of the binary logistic regression model greatly varies from both OL and OLS regression model. The reason for the difference is that the BL regression does not take into account the differences in positive malaria cases since it orders negative cases as 0 and all positive cases; children with yearly, monthly and quarterly cases as 1. Another reason is that 98% of the children with negative malaria cases are under 1 year of age. The BL model therefore does not take into account the differences in malaria occurrence for children above 1 to 5 years of age. Based on the limitation of the BL model for our analysis, we use OLS and OL regression results as the base for discussion.

Malaria and distance to health centre

Both OLS and OL results show a significant value between malaria and health centre. Interpreting the result with the estimate value, it can be explained that with an increase in the distance to health centre, the odds of having malaria is lower by 0.3863. This is similar to the findings of the correlated plot matrix which shows a negative correlation between the variables. This therefore implies that children living far away from the health centre are less likely to have malaria than those close by. This interpretation undermines the perception that people living close to the health centre are less likely to have malaria than those far away (Rutherford et al., 2009).

Beside the estimate interpretation, we see a correction in the result of the SEM in Appendix 5 which returned a non-significant value (p-value, 0.17697) between malaria and distance to health centre. Moreover, the SEM shows a significant pvalue for the Umugo ward (estimate= -0.0163147 and p-value= 0.00031), given with the same interpretation as explained earlier. The fact that the result explains less malaria occurrence for children far from health centre could as a result of the influence of external factors to health centre. Factors which may include access to vegetation where herbs used for malaria prevention can be easily accessed. Also, access to pharmacy "chemist" shop were malaria tablets can be obtained as pointed out in the findings of Kizito et al. (2012). These factors are uncovered in this analysis and thus need further validation.

Malaria and Room Occupancy Rate

These variables were found to be strongly correlated in both OLS and OL results, a finding similar to that derived from the scatter plot matrix. High pvalues were returned for both combined and individual ward analysis. Although in the OLS result, these variables were found to be insignificant for Umugo ward with estimate value=0.653603 and p-value=0.22789. Assessing the relationship between this variable and child malaria occurrence, it can be explained using the combined ward results that with an increase in room occupancy rate, the odds of having malaria increases by 0.5374. This means that children living in homes with more 2 persons per room are likely to have malaria more

frequently. This because a high sleeping density room favours the emanation of human odour that attracts mosquitoes indoors (Ngom & Siegmund, 2015).

Child malaria and distance to breeding site

The result of both OLS and OL regression show a significant relation between child malaria and distance to breeding sites; river and waste dump site. This is similar to major findings in malaria research such as: Rahman (2006) and Yusuf et al. (2010). The estimate value of these variables indicates that children living far away from these sites are less likely to have malaria than those living close by. However, there is a contrast between individual ward and the combined ward analysis. None of the individual wards have waste dumpsite and river significant.

For waste dumpsites, the differences in results could be linked to the lack of detailed data at an individual level. The field work conducted was unable to collect the location of waste dumpsite in the premises of individual households; this due to privacy reasons. For distance to the river, the differences in result could be linked to the fact that Nkpukpuevula; the ward close to the river is completely within 2,000metres from the river. Thus, all population are at risk of being bitten by mosquitoes. This could explain the reason for the non-variation, hence non-collinearity between malaria cases and distance to river in the sampled data. However, when comparing Nkpukpuevula to the other wards far from the river (Asa Amuhi and Umugo) as shown in the combined ward result, we see that river has a significant level of influence on malaria. This finding also validates the previous analyses done at the wardlevel which shows Nkpukpuevula to have highest number of malaria cases compared to other urban wards in Ugwunagbo area.

The differences in results, could also be linked to the influenced room occupancy rate and income has on the environmental variables, this as assessed from the correlation plot matrix in section 5.2.2.1. It could be that households with high room occupancy rate and low income level are more likely to live close to the river and waste dumpsite. Thus, there might not be a direct relation between river and malaria; relation might be influenced by the income and room occupancy rate. Since the relation between malaria and room occupancy rate as well as low income level is highly significant, a strong correlation between river and malaria could result. This assumption could be true, as it was observed during the field work, that Nkpukpuevula (the ward with highest number of malaria and at close distance to the river) is a slum; mainly occupied by low income households.

The OLS regression result shows water to be significant for Asa Amuhi (estimate=0.0056, pvalue=0.01533); an urban ward in Ugwunagbo. Since the actual distance was used for this regression (see: Table4.2), we interpret the estimate value as: an increase in the distance to water leads to a 0.0056 increase in the chances of having malaria. This means that children living far away from the "mapped" water surfaces are more likely to have malaria. This is unexpected as we expect children living close to water surfaces to have more malaria than those living far away. We therefore stick to the OL regression result which shows water to be insignificant to malaria at an individual level. The non-relation between malaria and

water could be linked to the strong association the "use of preventive health measure" variable has with malaria. This might also explain why waste dumpsite and river are not directly correlated with malaria based on the explanation made earlier (indirect influence of income level and occupancy rate on river and waste).

Child malaria and socio-economic condition

The relation between child malaria and socio-economic conditions; income level and educational level are highly significant in both OLS and OL results. This level of significance is present for the individual wards and for the combined ward analysis. The result is similar to the analysis derived from the scatter plot matrix and with the findings of Gayawan et.al(2014) and Yé et al (2006). Our result therefore validates the notion in malaria research that children from poor homes and with parents having low educational background are far more likely to have malaria than those from average and wealthy background.

Child malaria and the use of preventive health measures

The results of both OLS and OL results show a strong relation between malaria and the use of preventive health measures such as bed net. A significant level for both the individual wards and the combined ward result show these variables to be positively correlated. This finding validates the result of the correlation plot matrix which shows a very strong positive correlation between the assessed variables.

Child malaria and housing condition

The result of OL regression generally does not show a strong correlation between malaria and housing condition. Although the individual wards, Umugo and Asa Amuhi have wall and roof condition significant to malaria. The result for these wards is similar to that of the OLS regression in Appendix 5.

For window condition, there is no significance in the individual wards, although the aggregated ward result shows a strong relation. The discrepancy in the result for the assessed variables could be related to the significant level of influence preventive measure variable has on window condition; this as assessed in section 5.2.2.2 with the correlated plot matrix. It can be explained that children who use bed nets are more likely not to have their windows netted.

6. CONCLUSION AND RECOMMENDATION

The aim of this study was to use the existing frameworks in environmental inequality to define and model the relationship between urban malaria and environmental inequality. The model is tested using under 5 aged population of Aba South and Ugwunagbo areas. To actualize the aim of this research, 2 objectives were set: 1) to refine the existing frameworks in environmental inequality and link them to the factors of urban malaria to derive the relation between the two concepts 2) to identify correlation of urban childhood malaria and the inequality factors from the linked framework at different spatial aggregation level.

This research has answered the first objective by describing the causal factors of urban malaria in two groups: vector based factors (the environmental factors that enable the breed of mosquitoes) and the host based factors (the socio-economic and demographic factors that enable malaria transmission). The factors of environmental inequality has been described using the multi-level model of Kruize et al.,(2014). Factors are described at 3 aggregation levels: the macro level, the community level and the individual level. Using the multi-level model as a base model, the multi-differential malaria (MDM) model was derived. It aligned the factors of urban malaria with the 3 levels of environmental inequality. At the macro level are malaria risk factors such as: social, economic and political factors which lead to the distribution and creation of risk areas of malaria. At the community level are factors such as: housing quality and condition, crowding and mosquito breeding sites which lead to the differences in the level of exposure to malaria. At the individual level are malaria risk factors such as: age, income, educational level and use of clinical measures that lead to differences in health outcomes of malaria.

The second research objective was actualized using cluster analysis (done at a ward level) and regression analysis performed on the two levels (ward level and individual level) using the factors identified in the MDM model.

The result of the cluster analysis show wards (mainly: Aba Townhall and Mosque) at the north-eastern part of Aba South urban area to be hot spot locations of child malaria occurrence. These wards reoccurred as hot spot throughout the period 2013 – 2015.

The ward level regression analysis shows the number of houses within 1000 metres of the river to be correlated factors with malaria; although there were differences in the correlation between the areas. In Aba South urban area; where the hot spots were identified, river was found to be correlated with malaria and In Ugwunagbo urban area, water was found to be associated with malaria.

For the individual level analysis, three different wards were selected and household survey was conducted. A total of 300 samples were gathered; 100 for each selected ward. The result of the regression analysis done at this aggregate level show: the use of preventive health measures, income level, education level, room occupancy rate to be the most correlated factor to malaria occurrence.

The fact that distance to health centre was shown to be insignificant with child malaria could be linked to the small distances (sampled population to the nearest health centre) in our sample dataset. The research of Feikin et al.,(2009) and Larson et al.,(2012) used large distance ranges of 4 – 5 km for their analysis. Our study was unable to reach these distances as a result of the close distance the sampled wards are to health centre. Also, due to poor gps signal experienced during field work in the eastern part (2 km away from the health centre) of Asa Amuhi ward (See the map in figure 4.4 for details). To further test the relation between malaria and health centre, we recommend incorporating in the analysis, wards (such as: Nwaiyiekwe and Obegu) that are more than 3km away from the health centre.

Distance to water was found to be insignificant in the individual level analysis compared to the ward level result. This could be linked to the fact that our research lacked detailed information on puddles such as: the volume, the depth and the areas that could be flooded from the puddles. This information would have been used to model the flow accumulated areas from the puddles. This would have provided us with the actual water surface area that could have provided a good estimate in the individual regression results. The ward level analysis was able to detect this variable to be significant because the analytical method (number of households within threshold distance to water) used takes into account the areal extent that could be flooded from puddles.

The field work team were unable to collect waste dumpsites that are located in the premises of residential buildings. The mapped waste dumpsites therefore do not represent the actual number of waste dumpsites in the study area. This might have had an impact on the wardlevel regression result which shows waste dumpsites and malaria to be insignificant.

The analysis conducted in this research, points to the shortcoming of an analysis done at an aggregated level (ward level) and the risk in drawing inferences from this type of analysis. It can be seen from the combined ward analysis (using all the wards) in the ward level regression model that water was tested to be insignificant to child malaria occurrence. This same factor was shown to be significant in the individual analysis of Uguwanagbo urban area. Aggregated analysis, therefore does not model the actual correlated factor for all areas. It lacks details to the exact factors that are related to malaria in different areas. This research was able to identify through the individual analysis (done for each area) that water was significant to malaria in Ugwunagbo urban area while river was significant to Aba South area.

Individual level regression analysis compared to wardlevel regression analysis provides a more detailed result. It shows the underlining patterns unexplained at a ward level. This model was able to identify that the relation between malaria and river and waste dumpsite is based on the collinearity these variables has with room occupancy rate and income level.

The result of the cluster analysis and the exploratory analysis done in Chapter 3 shows that children in urban areas in the study area bear heavier burden of malaria than those in the

sub-urban and rural areas. Our findings is however, similar to the findings of Austin (2014); Fobil (2011) which proves urban areas to be more vulnerable to malaria due to income inequality, slum expansion and poor environmental condition. Our findings also validate the notion of Austin (2014) which shows that slum areas are more vulnerable to malaria than other areas. This was proven with the ward "Nkpukpuevula" which was identified as a slum area during the field work. The exploratory analysis shows this ward to have the highest number of cases compared to the other areas in Ugwungabo area. Our finding is however, contrary to the established theory in malaria research which proves rural areas to be more vulnerable to malaria than those in urban areas (Esimai & Aluko, 2015; Kigozi et al, 2015).

It could be seen from our research that the choice of regression model greatly affects regression result. The binary logistic (BL) regression model did not give a good estimate on the correlation between the assessed variables. The fact that it allows for only two categories (0 or 1) for the dependent variable made it difficult to retain the values and associated information of the different malaria case groups (no malaria, yearly, monthly, etc.) Also, because the negative malaria cases were mainly from a particular age group (less than 1 year), it was difficult to model effects for the rest of the aged groups. An alternative approach would have been to group children with yearly cases and those with negative cases as 0. The BL model should be suitable when negative and positive malaria cases vary across a study population.

The ordinal logistic (OL) regression model was a best fit for our data. The fact that it allows for multiple categories for the dependent variable made it possible to analyse the correlation between the various malaria case groups and the independent variables. Although researches suggest these two models (BL and OL) as an appropriate model for categorical variables (Lwetoijera et al., 2013 & Yusuf et al., 2010); our findings did not see any difference in the result of the Ordinary Least Square model and the OL model. Our research also points to the accuracy of spatial models. This can be seen in the correction it made on the results for distance to health centre variable which was shown to be significant in the results for OL and OLS models.

The univariate analysis (each independent variable) was applied because the result from the multi-variate regression analysis only had few variables (use of clinical measures and income level) significant. The result of this analysis shows multi-collinearity amongst the variables. Further tests were conducted using the variance factor index (Ndiath et al., 2015) to check if multi-collinearity poses any effect on the accuracy of the regression result. In this index, an independent variable with index value larger than 5 indicates that it is highly collinear and could affect regression result. Our result on this index showed all the independent variables with values less than 2.5 which suggest no effects.

The MDM model selected malaria factors that are identical to the inequality factors in multi-level model. Factors are streamlined to those factors that vary across space such as: income, neighbourhood conditions and population density. It therefore neglects environmental factors such as: altitude, temperature and rainfall which are significant factors to malaria and could vary at a larger geographic scale. This makes this framework only suitable for local

studies and not applicable for global studies. Moreover, the model is based on urban related malaria factors which might not be detailed enough to test malaria inequality in rural areas.

This research provides insight to the health department of Ugwunagbo and Aba South LGA on the areas of intervention for malaria. From the result of our analysis, we strongly recommend intervention in the following areas: Nkpukpuevula, Obeaja, Aba River, Aba Townhall, Igwebuike and Mosque. The distribution of treated bed nets should be concentrated on these areas to help alleviate the burden of malaria on children living these areas. Our research found the use of treated bed net to have a significant level of influence in the reduction of child malaria cases. It is therefore expected, that if the recommendation is followed, the desired outcomes will be reached.

REFERENCE

Abdulkadir, M. B., Ibraheem, R. M., & Johnson, W. B. R. (2015). Sociodemographic and Clinical Determinants of Time to Care-Seeking Among Febrile Children Under-Five in North-Central Nigeria. *Oman Medical Journal*, *30*(5), 331–5. http://doi.org/10.5001/omj.2015.68

Afon, A. (2012). A survey of operational characteristics, socioeconomic and health effects of scavenging activity in Lagos, Nigeria. *Waste Management & Research : The Journal of the International Solid Wastes and Public Cleansing Association, ISWA*, *30*(7), 664–71. http://doi.org/10.1177/0734242X12444894

Afrane, Y. A., Lawson, B. W., Brenya, R., Kruppa, T., & Yan, G. (2012). The ecology of mosquitoes in an irrigated vegetable farm in Kumasi, Ghana: abundance, productivity and survivorship. *Parasites & Vectors*, *5*(1), 233. http://doi.org/10.1186/1756-3305-5-233

Akono, P. N., Mbida, J. A. M., Tonga, C., Belong, P., Ngo Hondt, O. E., Magne, G. T., … Lehman, L. G. (2015). Impact of vegetable crop agriculture on anopheline agressivity and malaria transmission in urban and less urbanized settings of the South region of Cameroon. *Parasites & Vectors*, *8*(1), 293. http://doi.org/10.1186/s13071-015-0906-2

Anselin, L. (2005). Exploring Spatial Data with GeoDa: A Workbook. *Geography*, 244. http://doi.org/http://www.csiss.org/

Austin, K. F. (2014). Dependency, Urban Slums, and the Forgotten Plagues: Tuberculosis and Malaria Prevalence in Less Developed Nations. *Sociological Perspectives*, *58*(2), 286–310. http://doi.org/10.1177/0731121414556542

Banerjee, D. (2013). Understanding Environmental Inequalities: A Cross-Disciplinary Review of Research Approaches. *Environmental Justice*, *6*(4), 140–144. http://doi.org/10.1089/env.2013.0009

Brulle, R. J., & Pellow, D. N. (2006). Environmental justice: human health and environmental inequalities. *Annual Review of Public Health*, *27*, 103–24. http://doi.org/10.1146/annurev.publhealth.27.021405.102124

CHAN., Yuan Hwa CHOU., Wei HAO., R. E. (2014). BIOSTATISTICS IN PSYCHIATRY, *Vol.26*, 105 – 109.

Craig, M. H., Snow, R. W., & le Sueur, D. (1999). A Climate-based Distribution Model of Malaria Transmission in Sub-Saharan Africa. *Parasitology Today*, *15*(3), 105–111. http://doi.org/10.1016/S0169-4758(99)01396-4

De Silva, P. M., & Marshall, J. M. (2012). Factors contributing to urban malaria transmission in sub-saharan Africa: a systematic review. *Journal of Tropical Medicine*, *2012*, 819563. http://doi.org/10.1155/2012/819563

Deguen, S., & Zmirou-Navier, D. (2010). Social inequalities resulting from health risks related to ambient air quality--A European review. *European Journal of Public Health*, *20*(1), 27–35. http://doi.org/10.1093/eurpub/ckp220

Demographic Household Survey(DHS). (2013). *Household Population and Housing Characteristics*. Retrieved from https://dhsprogram.com/pubs/pdf/FR293/FR293.pdf

Dennis, R., Howick, R., Stewart, N., Bellamy, G., & Fairclough House, R. (n.d.). Science Report Methods of Estimating Population and Household Projections 2. Retrieved from www.environment-agency.gov.uk

Dinda, S. (2004). Environmental Kuznets Curve Hypothesis: A Survey. *Ecological Economics*, *49*(4), 431–455. http://doi.org/10.1016/j.ecolecon.2004.02.011

Donnelly, M. J., McCall, P. J., Lengeler, C., Bates, I., D'Alessandro, U., Barnish, G., … Mutero, C. (2005). Malaria and urbanization in sub-Saharan Africa. *Malaria Journal*, *4*, 12. http://doi.org/10.1186/1475-2875-4-12

Ernst, K. C., Adoka, S. O., Kowuor, D. O., Wilson, M. L., & John, C. C. (2006). Malaria hotspot areas in a highland Kenya site are consistent in epidemic and non-epidemic years and are associated with ecological factors. *Malaria Journal*, *5*, 78. http://doi.org/10.1186/1475-2875-5-78

Esimai, O. A., & Aluko, O. O. (2015). Determinants of use of insecticide treated bednets among caregivers of under five children in an urban local government area of Osun state, South-Western Nigeria. *Global Journal of Health Science*, *7*(2), 20–7. http://doi.org/10.5539/gjhs.v7n2p20

Evans, G. W., & Kantrowitz, E. (2002). Socioeconomic status and health: the potential role of environmental risk exposure. *Annual Review of Public Health*, *23*, 303–31. http://doi.org/10.1146/annurev.publhealth.23.112001.112349

Feikin, D. R., Nguyen, L. M., Adazu, K., Ombok, M., Audi, A., Slutsker, L., & Lindblade, K. A. (2009). The impact of distance of residence from a peripheral health facility on pediatric health utilisation in rural western Kenya. *Tropical Medicine & International Health*, *14*(1), 54–61. http://doi.org/10.1111/j.1365-3156.2008.02193.x

Fobil, J. N., Kraemer, A., Meyer, C. G., & May, J. (2011). Neighborhood urban environmental quality conditions are likely to drive malaria and diarrhea mortality in Accra, Ghana. *Journal of Environmental and Public Health*, *2011*, 484010. http://doi.org/10.1155/2011/484010

Gayawan, E., Arogundade, E. D., & Adebayo, S. B. (2014). A Bayesian multinomial modeling of spatial pattern of co-morbidity of malaria and non-malarial febrile illness among young children in Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *108*(7), 415–24. http://doi.org/10.1093/trstmh/tru068

Hamilton, A. J., Burry, K., Mok, H.-F., Barker, S. F., Grove, J. R., & Williamson, V. G. (2013). Give peas a chance? Urban agriculture in developing countries. A review. *Agronomy for Sustainable Development*, *34*(1), 45–73. http://doi.org/10.1007/s13593-013-0155-8

Kaika, D., & Zervas, E. (2013). The Environmental Kuznets Curve (EKC) theory—Part A: Concept, causes and the CO2 emissions case. *Energy Policy*, *62*, 1392–1402. http://doi.org/10.1016/j.enpol.2013.07.131

Kaufmann, C., & Briegel, H. (2004). Flight performance of the malaria vectors Anopheles gambiae and Anopheles atroparvus.

Kigozi, S. P., Pindolia, D. K., Smith, D. L., Arinaitwe, E., Katureebe, A., Kilama, M., … Tatem, A. J. (2015). Associations between urbanicity and malaria at local scales in Uganda. *Malaria Journal*, *14*(1), 374. http://doi.org/10.1186/s12936-015-0865-2

Kizito, J., Kayendeke, M., Nabirye, C., Staedke, S. G., & Chandler, C. I. (2012). Improving access to health care for malaria in Africa: a review of literature on what attracts patients. *Malaria Journal*, *11*(1), 55. http://doi.org/10.1186/1475-2875-11-55

Klinkenberg, E., McCall, P., Wilson, M. D., Amerasinghe, F. P., & Donnelly, M. J. (2008). Impact of urban agriculture on malaria vectors in Accra, Ghana. *Malaria Journal*, *7*, 151. http://doi.org/10.1186/1475-2875-7-151

Konradsen, F., Amerasinghe, P., Van Der Hoek, W., Amerasinghe, F., Perera, D., & Piyaratne, M. (2003). Strong association between house characteristics and malaria vectors in Sri Lanka. *American Journal of Tropical Medicine and Hygiene*, *68*(2), 177–181. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0037316016&partnerID=tZOtx3y1

Kruize, H. (2007). On environmental equity : Exploring the distribution of environmental quality among socio-economic categories in the Netherlands. Retrieved from http://www.narcis.nl/publication/RecordID/oai%3Adspace.library.uu.nl%3A1874%2F22609

Kruize, H., Droomers, M., van Kamp, I., & Ruijsbroek, A. (2014). What causes environmental inequalities and related health effects? An analysis of evolving concepts. *International Journal of Environmental Research and Public Health*, *11*(6), 5807–27. http://doi.org/10.3390/ijerph110605807

Larson, P. S., Mathanga, D. P., Campbell, C. H., & Wilson, M. L. (2012). Distance to health services influences insecticide-treated net possession and use among six to 59 month-old children in Malawi. *Malaria Journal*, *11*(1), 18. http://doi.org/10.1186/1475-2875-11-18

Leandro-Reguillo, P., Thomson-Luque, R., Monteiro, W. M., & de Lacerda, M. V. G. (2015). Urban and architectural risk factors for malaria in indigenous Amazonian settlements in Brazil: a typological analysis. *Malaria Journal*, *14*(1), 284. http://doi.org/10.1186/s12936-015-0806-0

Liu, L. (2012). Environmental poverty, a decomposed environmental Kuznets curve, and alternatives: Sustainability lessons from China. *Ecological Economics*, *73*, 86–92. http://doi.org/10.1016/j.ecolecon.2011.10.025

Lwetoijera, D. W., Kiware, S. S., Mageni, Z. D., Dongus, S., Harris, C., Devine, G. J., & Majambere, S. (2013). A need for better housing to further reduce indoor malaria transmission in areas with high bed net coverage. *Parasites & Vectors*, *6*, 57. http://doi.org/10.1186/1756-3305-6-57

Maantay, J. (2002). Mapping environmental injustices: Pitfalls and potential of geographic information systems in assessing environmental health and equity. *Environmental Health Perspectives*. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0036227366&partnerID=tZOtx3y1

Marquetti, M. del C., Rojas, L., Mohd Birniwa, M., Sulaiman, H. U., & Adamu, H. H. (2007).

Identificación de los sitios de cría de Anopheles sp. durante parte de la estación seca en el estado de Jigawa, Nigeria. *Revista Cubana de Medicina Tropical*, *59*(2), 166–168. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-77953736327&partnerID=tZOtx3y1

MARTENS, P., KOVATS, R., NIJHOF, S., DEVRIES, P., LIVERMORE, M., BRADLEY, D., … MCMICHAEL, A. (1999). Climate change and future populations at risk of malaria. *Global Environmental Change*, *9*, S89–S107. http://doi.org/10.1016/S0959-3780(99)00020-5

Mckenna, S., Meyer, M., Gregg, C., & Gerber, S. (2015). s-CorrPlot: An Interactive Scatterplot for Exploring Correlation. Retrieved from http://sci.utah.edu/~vdl/papers/2015_jcgs_s-corrplot.pdf

Morello-Frosch, R., & Lopez, R. (2006). The riskscape and the color line: examining the role of segregation in environmental health disparities. *Environmental Research*, *102*(2), 181–96. http://doi.org/10.1016/j.envres.2006.05.007

National Population Commission. (2010). Nigeria Malaria Indicator Survey (MIS).

Ndiath, M. M., Cisse, B., Ndiaye, J. L., Gomis, J. F., Bathiery, O., Dia, A. T., … Faye, B. (2015). Application of geographically-weighted regression analysis to assess risk factors for malaria hotspots in Keur Soce health and demographic surveillance site. *Malaria Journal*, *14*(1), 463. http://doi.org/10.1186/s12936-015-0976-9

Ngom, R., & Siegmund, A. (2015). The key role of socio-demographic and socio-environmental factors in urban malaria occurrence and control - An illustration using the city of Yaoundé. *Social Science & Medicine (1982)*, *133*, 269–79. http://doi.org/10.1016/j.socscimed.2014.12.001

Nriagu, J., Afeiche, M., Linder, A., Arowolo, T., Ana, G., Sridhar, M. K. C., … Adesina, A. (2008). Lead poisoning associated with malaria in children of urban areas of Nigeria. *International Journal of Hygiene and Environmental Health*, *211*(5-6), 591–605. http://doi.org/10.1016/j.ijheh.2008.05.001

O'Neill, M. S., Jerrett, M., Kawachi, I., Levy, J. I., Cohen, A. J., Gouveia, N., … Zanobetti, A. (2003). Health, wealth, and air pollution: Advancing theory and methods. *Environmental Health Perspectives*. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-10744225114&partnerID=tZOtx3y1

Oduola, A. O., Olojede, J. B., Oyewole, I. O., Otubanjo, O. A., & Awolola, T. S. (2013). Abundance and diversity of Anopheles species (Diptera: Culicidae) associated with malaria transmission in human dwellings in rural and urban communities in Oyo State, Southwestern Nigeria. *Parasitology Research*, *112*(10), 3433–9. http://doi.org/10.1007/s00436-013-3522-0

Ogunrinola, I. O., & Adepegba, E. O. (2012). Health and Economic Implications of Waste Dumpsites in Cities: The Case of Lagos, Nigeria. *International Journal of Economics and Finance*. Retrieved from http://ccsenet.org/journal/index.php/ijef/article/view/15929

Okezie, C. ., Sulaiman, J., & Nwosu, A. . (2012). Farm – Level Determinants of Agricultural Commercialization. Retrieved from

http://article.sapub.org/10.5923.j.ijaf.20120202.01.html

Pellow, D. N. (2000). Environmental inequality formation: Toward a theory of environmental injustice. *American Behavioral Scientist*, (4), 581–601. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0033634148&partnerID=tZOtx3y1

PELLOW, D. N. (2000). Environmental Inequality Formation: Toward a Theory of Environmental Injustice. *American Behavioral Scientist*, *43*(4), 581–601. http://doi.org/10.1177/0002764200043004004

Petersen, E., Severini, C., & Picot, S. (2013). Plasmodium vivax malaria: a re-emerging threat for temperate climate zones? *Travel Medicine and Infectious Disease*, *11*(1), 51–9. http://doi.org/10.1016/j.tmaid.2013.01.003

Peterson, I., Borrell, L. N., El-Sadr, W., & Teklehaimanot, A. (2009). A temporal-spatial analysis of malaria transmission in Adama, Ethiopia. *The American Journal of Tropical Medicine and Hygiene*, *81*(6), 944–9. http://doi.org/10.4269/ajtmh.2009.08-0662

Ping, J. L., Green, C. J., Zartman, R. E., & Bronson, K. F. (2004). Exploring spatial dependence of cotton yield using global and local autocorrelation statistics, 220 – 236. http://doi.org/10.1016/j.fcr.2004.02.009

Rahman, A. (2006). Assessing income-wise household environmental conditions and disease profile in urban areas: Study of an Indian city. *GeoJournal*, *65*(3), 211–227. http://doi.org/10.1007/s10708-005-3127-y

Ranjit Kumar. (2014). *Research Methodology, A Step-by-Step Guide For Beginners* (4th ed.). London: Sage. Retrieved from https://wordery.com/research-methodology-ranjit-kumar-9781446269978?currency=GBP&gtrck=MUcvQlcrbDM4bFZoQnpzTHA1VzRaZXUrVEhsbG96ZjdvMURRqWJcwVHprOCtrRzkzaXptenhlYXdqeXIyM2V0UEZqdWMxcHVzeDhJS2ZGNmdkaXRLYUE9PQ&gclid=CJrO8fDe1s4CFRaNGwod8GoOFw

Robert, V., Macintyre, K., Keating, J., Trape, J. F., Duchemin, J. B., Warren, M., & Beier, J. C. (2003). Malaria transmission in urban sub-Saharan Africa. *American Journal of Tropical Medicine and Hygiene*. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0037315603&partnerID=tZOtx3y1

Rosas-Aguirre, A., Ponce, O. J., Carrasco-Escobar, G., Speybroeck, N., Contreras-Mancilla, J., Gamboa, D., … Llanos-Cuentas, A. (2015). Plasmodium vivax malaria at households: spatial clustering and risk factors in a low endemicity urban area of the northwestern Peruvian coast. *Malaria Journal*, *14*(1), 176. http://doi.org/10.1186/s12936-015-0670-y

Rutherford, M. E., Dockerty, J. D., Jasseh, M., Howie, S. R. C., Herbison, P., Jeffries, D. J., … Hill, P. C. (2009). Access to health care and mortality of children under 5 years of age in the Gambia: a case-control study. *Bulletin of the World Health Organization*, *87*(3), 216–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19377718

Rutherford, M. E., Mulholland, K., & Hill, P. C. (2010). How access to health care relates to under-five mortality in sub-Saharan Africa: systematic review. *Tropical Medicine & International Health*, *15*(5), 508–519. http://doi.org/10.1111/j.1365-

3156.2010.02497.x

Sedgwick, P. (2012). Pearson's correlation coefficient. http://doi.org/10.1136/bmj.e4483

Soobader, M., Cubbin, C., Gee, G. C., Rosenbaum, A., & Laurenson, J. (2006). Levels of analysis for the study of environmental health disparities. *Environmental Research*, *102*(2), 172–80. http://doi.org/10.1016/j.envres.2006.05.001

Szasz, A., & Meuser, M. (2000). Unintended, inexorable: The production of environmental inequalities in Santa Clara County, California. *American Behavioral Scientist*, (4), 602–632. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0033632517&partnerID=tZOtx3y1

Tanser, F. C., Sharp, B., & le Sueur, D. (2003). Potential effect of climate change on malaria transmission in Africa. *Lancet (London, England)*, *362*(9398), 1792–8. http://doi.org/10.1016/S0140-6736(03)14898-2

Umeh, J. M. (2013). *a Hospital Based Study of Malaria in Ndiegoro …*. Retrieved from http://www.bing.com/search?q=The+prevalence+of+Malaria+in+Aba%2C+Abia+State&qs=n&form=QBRE&pq=the+prevalence+of+malaria+in+aba%2C+abia+state&sc=0-0&sp=-1&sk=&cvid=d7e23919e49741e7bda1dd051c33f03f

Wachukwu, K. C., & Eleanya, E. U. (2007). Health impact assessment of solid waste disposal workers in Port Harocurt, Nigeria. *Journal of Applied Sciences*, *7*(22), 3562–3566. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-36849087324&partnerID=tZOtx3y1

World Health Organisation Europe. (2010). WHO, Europe; Environmental health inequalities in Europe. Retrieved February 8, 2016, from http://www.euro.who.int/__data/assets/pdf_file/0010/157969/e96194.pdf

World Health Organisation(WHO). (2015). World Malaria Report 2015. Retrieved from http://apps.who.int/iris/bitstream/10665/200018/1/9789241565158_eng.pdf

Wright, R. J., & Subramanian, S. V. (2007). Advancing a multilevel framework for epidemiologic research on asthma disparities. *Chest*, *132*(5 Suppl), 757S–769S. http://doi.org/10.1378/chest.07-1904

Ye, F. (2013). Comparing Three Commonly Used Crash Severity Models on Sample Size Requirements: Multinomial Logit, Ordered Probit and Mixed Logit Models.

Yé, Y., Hoshen, M., Louis, V., Séraphin, S., Traoré, I., & Sauerborn, R. (2006). Housing conditions and Plasmodium falciparum infection: protective effect of iron-sheet roofed houses. *Malaria Journal*, *5*, 8. http://doi.org/10.1186/1475-2875-5-8

Yusuf, O. B., Adeoye, B. W., Oladepo, O. O., Peters, D. H., & Bishai, D. (2010). Poverty and fever vulnerability in Nigeria: a multilevel analysis. *Malaria Journal*, *9*(1), 235. http://doi.org/10.1186/1475-2875-9-235

Zhang, Y., & Maciejewski, R. (2017). Quantifying the Visual Impact of Classification Boundaries in Choropleth Maps. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 371–380. http://doi.org/10.1109/TVCG.2016.2598541

APPENDIX

Appendix 1: Questionnaire

PERSONAL DETAILS

A. House serial number and ward location ……………………………………………………………….

B. Spatial reference of residence, latitude……….. ……… and longitude…………………………………

C. Age of child <1 ☐1 ☐2 ☐ ☐ ☐5 ☐

D. Gender ☐ Male ☐ Female

E. Occupation of parent(s): Farmer ☐ Teacher☐ Civil servant☐ Trader ☐ Unemployed☐
Others Specify ………………………………………………………………….

F. Income level per month
<20,000 ☐ 21,000 – 40,000☐ 41,000 – 60,000 ☐ 61,000 and above ☐

G. Educational level of parents: First school leaving and below☐ technical and commercial
☐WAEC☐NCE/OND☐ HND/BSC ☐ MSC☐ PHD ☐ Others specify…………………………

HOUSING DETAILS

H. Number of family members …………………….. number of rooms………………………….

I. Type of housing material (roof, wall)
mud wall☐ brick/ cement wall ☐ tile wall☐ others specify ……………………………………

thatched roof☐ zinc roof☐ asbestos roof ☐others specify …………………………………

K. Condition of housing unit (roof, walls and windows)
Cracked walls ☐uncracked walls ☐
Opening in roof ☐ closed roof☐
Unscreened windows☐netted (screened) windows☐

HEALTH (MALARIA) DETAILS

L. When was the last time your child had malaria <4 weeks ago ☐ 2-3months ago ☐
4-6 months ago ☐ others specify……………………………………………………………………

M. How often does your child have malaria weekly ☐ monthly☐ yearly☐

N. What health care service did you (or do you visit) in times of illness pharmacy ☐
clinic/hospital☐, please specify name and location…………………………………

O.What preventive measures do you used to protect your child from malaria
   Treated net ☐   Insecticide(shelltox) ☐   Coil ☐   Traditional herbs☐
   pre-malariadrug☐   please   specify   how   often   you   give
   this…………………………………………..

P. How long does your child play outside daily, specify a duration(in
minutes/hours)……………………………………………….

APPENDIX 2: IMAGE OF PUDDLES AND WASTE DUMP SITE
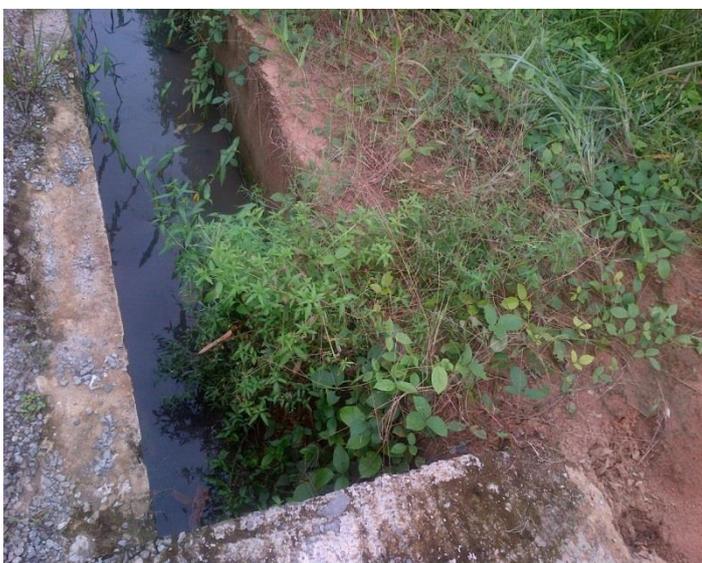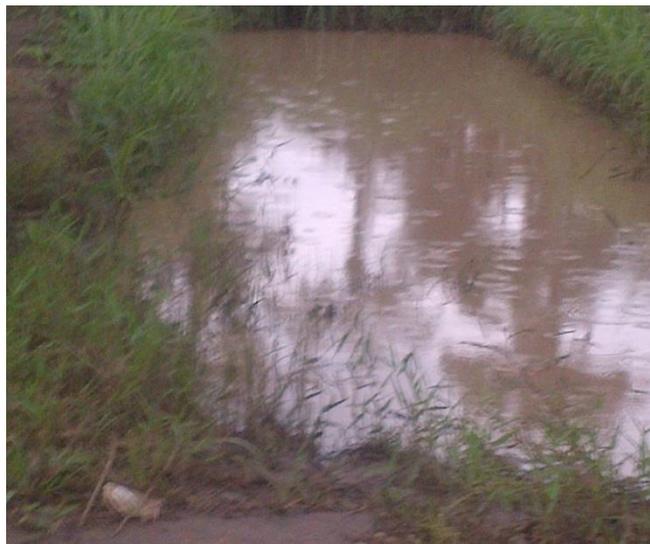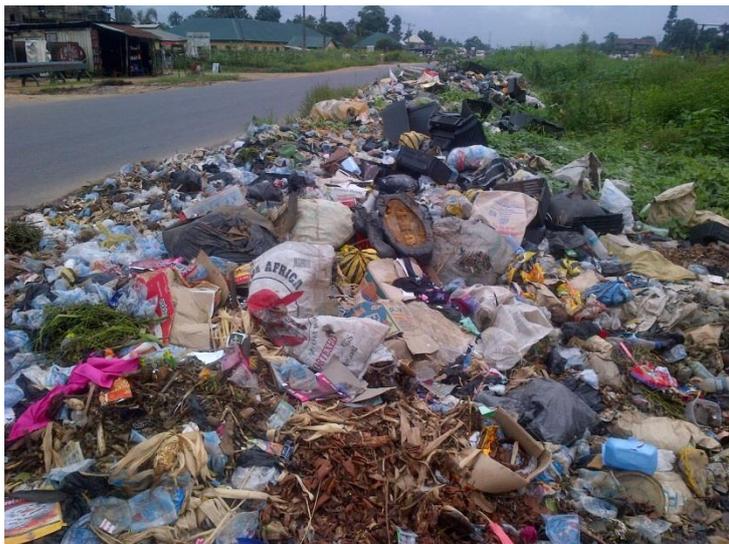
Appendix 3: Graphical representation of Child malaria cases per 1000 population.

Child malaria cases per 1000 population, 2013

0	Asa
0	Ekeoha
0	Glocester
4.4	Iheorji
5.1	Asa Amuhi
8.2	Ngwa
8.4	Umuchima
8.9	Umuarukwu
9.9	Ihie Obeaku
10.7	Umuada
11.6	Aba River
14	Okporoenyi
15.4	Umuogele
21.6	Owerri Aba
22.9	Eluohazu
24.5	Abayi Mbasaa
35.7	Amapu Ideobia
41.2	Asa Umunka

Child malaria cases per 1000 population, 2014

0	Asa
0	Ekeoha
2.9	Iheorji
4.3	Asa Amuhi
7.5	Umuchima
8.4	Umuada
8.8	Ihie Obeaku
21.3	Umuarukwu
28.6	Abayi Mbasaa
29	Owerri Aba
29.4	Amapu Ideobia
31.7	Ngwa
40.4	Umuogele
41.9	Obegu
47.8	Asa Umunka
52	Umugo
53.3	Glocester
73.3	Enyimba

Child Malaria cases per 1000, 2015

0	Asa
0	Ekeoha
0	Asa Amuhi
2.3	Ihie Obeaku
13.1	Ngwa
13.6	Umuada
19.1	Okporoenyi
22.9	Iheorji
24	Umuarukwu
24.8	Abayi Mbasaa
25.1	Glocester
28.7	Umuchima
32.9	Enyimba
33.4	Umuogele
49.3	Ihie Ukwu
49.6	Asa Umunka
54.5	Owerri Aba
67.4	Eluohazu

Appendix 4: Spatial Dependency Test for Spatial Model Selection

TEST	MI/DF	VALUE	PROB
Moran's I (error)	-0.0548	-1.6430	0.10038
Lagrange Multiplier (lag)	1	0.0831	0.77311
Robust LM (lag)	1	0.8290	0.36256
Lagrange Multiplier (error)	1	1.#INF	0.00000
Robust LM (error)	1	1.#INF	0.00000
Lagrange Multiplier (SARMA)	2	1.#INF	0.00000

Appendix 5: OLS and SEM Regression Results for Individual level Analysis

Variables	All wards		Asa Amuhi		Umugo		Nkpukpuevula	
	estimate	pvalue	estimate	pvalue	estimate	pvalue	estimate	pvalue
Housing condition(OLS)	0.9906	0.00673**	0.4398	0.00005***	0.5794	0.05742	-0.4584	0.54457
Housing condition (SEM)	0.923823	0.00486**	2.13454	0.00000***	1.14325	0.05265	-0.264985	0.75130
Occupancy rate (OLS)	1.6053	0.00000***	1.637	0.00227**	0.653603	0.22789	0.7790	0.03179*
Occupancy rate (SEM)	1.11808	0.00001***	1.73907	0.00071***	0.822199	0.21616	1.1736	0.00311**
Preventive measure (OLS)	2.7289	0.00000***	0.2989	0.00000***	2.3317	0.00000***	2.4029	0.00000***
Preventive measure (SEM)	2.62896	0.00000***	2.7124	0.00000***	2.48888	0.00000***	2.47116	0.00000***
Income level (OLS)	3.151	0.00000***	0.7050	0.00906***	2.1520	0.05197*	3.2564	0.01883*
Income level (SEM)	2.37995	0.00002***	2.01387	0.00384**	2.47929	0.03021*	4.75548	0.00206**
Education level (OLS)	4.4756	0.00000***	1.014	0.00033***	2.3850	0.03066*	3.6095	0.00004***
Educational level (SEM)	4.07951	0.00000***	4.5478	0.00001***	2.58391	0.01856*	3.90264	0.00001***
Distance to health centre (OLS)	-0.0029	0.01130*	0.0057	0.01080*	-0.0001	0.93199	-0.0044	0.16740
Distance to health centre (SEM)	-0.0032	0.17697	-0.0163147	0.00031***	-0.000158255	0.93140	-0.00475088	0.25330
Distance to river (OLS)	-0.0005	0.00000***	0.0014	0.70332	0.0002	0.68798	-0.0000	0.99306

Distance to river (SEM)	-0.0006	0.00004***	-0.0005	0.78215	0.000283965	0.71346	0.0000	0.99861
Distance to waste (OLS)	-0.0013	0.03296*	0.0008	0.69426	0.0003	0.79932	-0.00142655	0.63599
Distance to waste (SEM)	-0.0025	0.03001*	0.00034699	0.75411	0.00037096	0.80069	-0.00140012	0.70596
Distance to water (OLS)	-0.0009	0.45142	0.0056	0.01533*	0.0004	0.79706	-0.00038345	0.90486
Distance to water (SEM)	-0.0037	0.14256	0.112677	0.01146*	0.00040809	0.79953	-0.000345738	0.93042

Appendix 6: Binary Logistic Regression Result

Variable	All wards		
	Estimate	STD.error	Pvalue
Distance to the nearest health center	0.0002109	0.0005809	0.717
Distance to the river	0.0000	0.0000	0.123
Distance to the nearest water surface	0.0002624	0.0006321	0.678
Distance to the nearest waste dump site	-0.0004835	0.0002836	0.0882
Occupancy rate	0.3655	0.1638	0.0257 *
Wall condition	0.07201	0.51219	0.888
Roof condition	-0.1345	0.4513	0.766
Window condition	-0.07746	0.36013	0.83

Appendix 7: Python Script Created in "R" Spatial Statistical Package

```
##Descriptive Statistics and Regression Analysis of Child Malaria, WardLevel Analysis
##MSC Thesis
##created by Chibunna Ann Orji
## Created 10th of October, 2016 and Edited on Febuary 10, 2017
##Descriptive statistics Wardlevel Analysis
#read data into variable
wardreg<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\allwardv03.csv"),header=T,sep=",")

 #display all data
 wardreg
 summary(wardreg)
 attach(wardreg)

 ##BOX PLOT

 #boxplot of dependent variables
 boxplot(Malaria_14,Malaria_15)
 boxplot(Malaria_14, Malaria_15, ylab= "malariaper1000", xlab= "2014 and 2015", main="Boxplot Distribution of
Childhood Malaria", col='red')
```

```r
#boxplot of explanatory variables
boxplot(Pop_den,water,river,waste, ylab="Number of houses within threshold distance/Popdensity",
xlab="Pop_den,Water,River,Waste", main="Boxplot distribution of the explanatory variables", col='red')


##TEST FOR NORMAL DISTRIBUTION USING HISTOGRAM

#histogram of dependent variable Malaria2014
hist( Malaria_14, xlab="Malaria per 1000 population", main="Histogram Malaria Incident rate 2014", col= "blue")
#histogram of dependent variable Malaria2015
hist( Malaria_15, xlab="Malaria per 1000 population", main="Histogram Malaria Incident rate 2015", col= "blue")
#histogram of dependent variable Population density
hist( pop_den, xlab="Population density", main="Histogram of Population density", col= "blue")
#histogram of dependent variable Water surfaces
hist( water, xlab="Distance to water surfaces",main="Histogram of the Distance to Water Variable", col= "blue")
#histogram of dependent variable River
hist( river, xlab="Distance to river",main="Histogram of the Distance to River Variable", col= "blue")
#histogram of dependent variable Waste
hist( waste, xlab="Distance to Waste site",main="Histogram of the Distance to Waste Variable", col= "blue")
#histogram of dependent variable Health centre
hist( h.c, xlab="Distance to health centre",main="Histogram of the Distance to Health Centre Variable", col=
"blue")
#histogram of explanatory variable population density
hist(suburban$pop_den, xlab="population density", main="Histogram of Population density")
#histogram of explanatory variable Artificial water surfaces
hist(suburban$Water, xlab="Houses within 200m to water surfaces", main="Histogram of the distance to water
surfaces")
#histogram of explanatory variable Waste dumpsite
hist(suburban$Waste, xlab="Houses within 200m to waste dump", main="Histogram of the distance to wate
dumpsite")


#CORRELATION ANALYSIS USING SCATTER PLOT AND LOG TRANSFORMED DATA

#Correlation matrix for all wards
#read data into variable
matAB<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\regressionperward\\lo
g_allwardv03.csv"),header=T,sep=",")
#display all data
matAB
attach(matAB)
library(corrplot)
matBD<-cor(matAB)
corrplot(matBD,order="hclust", addrect=2)


#Correlation matrix Ugwunagbo Urban

#read data into variable
matCB<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\regressionperward\\H
C_Ugwuregression\\Log_UgwuUrban.csv"),header=T,sep=",")
#display all data
matCB
attach(matCB)
library(corrplot)
matDA<-cor(matCB)
corrplot(matDA,order="hclust", addrect=2)
```

```
#Correlation matrix Ugwunagbo Rural

#read data into variable

matCD<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\regressionperward\\H
C_Ugwuregression\\Log_UgwuRural.csv"),header=T,sep=",")

#display all data
matCD
attach(matCD)
library(corrplot)
matDB<-cor(matCD)
corrplot(matDB,order="hclust", addrect=2)
 plot(LogMalaria14,LogWater)


###Correlation plot for AbaSouth
###Correlation matrix Ugwunagbo Rural and health centre 2014 and 2015
#read data into variable
matCE<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\regressionperward\\Lo
g_Abasouthv02.csv"),header=T,sep=",")

#display all data
matCE
attach(matCE)
library(corrplot)
matDC<-cor(matCE)
corrplot(matDC,order="hclust", addrect=2)


##OLS REGRESSION CHILD MALARIA 2014/2015

##all wards including abasouth and ugwunagbo

attach(wardco)
wardco
reg_all<-lm(log_14 ~ log_water + log_waste + Pop_den + log_river)
summary(reg_all)

reg_all_15<-lm(log_15 ~ log_water + log_waste + Pop_den + log_river)
summary(reg_all_15)


##urban wards in aba south

#read data into variable

reg_absouth<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\regressionperwa
rd\\Log_Abasouth.csv"),header=T,sep=",")

reg_absouth
attach(reg_absouth)
reg_abs<-lm(log_14 ~ log_water + log_waste + Pop_den + log_river, data=reg_absouth)
summary(reg_abs)
reg_abs_15<-lm(log_15 ~ log_water + log_waste + Pop_den + log_river, data=reg_absouth)
summary(reg_abs_15)
```

```
##regression urban wards in Ugwunagbo

#read data into variable

reg_ugw_ub<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\regressionperwa
rd\\Log_Ugwu_urban.csv"),header=T,sep=",")

reg_ugw_ub
attach(reg_ugw_ub)
reg_ub<-lm(log_14 ~ log_water + log_waste + Pop_den + log_river + h.c_catg, data=reg_ugw_ub)
summary(reg_ub)
reg_ub2<-lm(log_14 ~ log_river)
summary(reg_ub2)
reg_ub1<-lm(log_15 ~ log_water + log_waste + Pop_den + log_river + h.c_catg, data=reg_ugw_ub)
summary(reg_ub1)


##rural wards in ugwunagbo

#read data into variable

reg_ugw_ru<read.table(("C:\\Users\\Anna\\Documents\\final_regression\\Allwardregression\\regressionperwar
d\\Log_Ugwurural.csv"),header=T,sep=",")

reg_ugw_ru
attach(reg_ugw_ru)
reg_ur<-lm(log_14 ~ log_water + log_waste + Pop_den + log_river + h.c_catg, data=reg_ugw_ru)
summary(reg_ur)
reg_ur1<-lm(log_15 ~ log_water + log_waste + Pop_den + log_river + h.c_catg, data=reg_ugw_ru)
summary(reg_ur1)
reg_ur2<-lm(log_15 ~ log_water, data=reg_ugwu_ru)
summary(reg_ur2)
reg_ur2<-lm(log_15 ~ log_water, data=reg_ugwu_ru)
summary(reg_ur2)
reg_ur2<-lm(log_15 ~ log_water, data=reg_ugwu_ru)
summary(reg_ur2)
reg_ur2<-lm(log_15 ~ log_water, data=reg_ugwu_ru)
summary(reg_ur2)
```