Neighbourhood-level housing quality indices for health assessment in Dakar, Senegal

Assane N. Gadiaga,[1,2] Florence De Longueville,[1,2] Stefanos Georganos,[3] Tais Grippa,[3] Sébastien Dujardin,[1,2] Aminata Niang Diène,[4] Bruno Masquelier,[5] Mouhamadou Diallo,[6] Catherine Linard[1,2]

[1]Institute of Life, Earth and Environment, Université de Namur, Namur, Belgium; [2]Department of Geography, Université de Namur, Namur, Belgium; [3]Department of Geoscience, Environment & Society, Université Libre De Bruxelles, Bruxelles, Belgium; [4]Department of Geography, Université Cheikh Anta Diop, Dakar, Senegal; [5]Center for Demographic Research, Université catholique de Louvain, Louvain-la-Neuve, Belgium; [6]Molecular Biology Unit/Bacteriology-virology laboratory, Université Cheikh Anta Diop, Dakar, Senegal

Abstract

In sub-Saharan African cities, the dearth of accurate and detailed data is a major problem in the study of health and socio-economic changes driven by rapid urbanization. Data on both health determinants and health outcomes are often lacking or are of poor quality. Proxies associated with socioeconomic differences are needed to compensate the lack of data. One of the most straightforward proxies is housing quality, which is a multidimensional concept including characteristics of both the built and natural environments. In this work, we combined the 2013 census data with remotely sensed land cover and land use data at a very high resolution in order to develop an integrated housing quality-based typology of the neighbourhoods in Dakar, Senegal. Principal component analysis and hierarchical classification were used to derive neighbourhood housing quality indices and four neighbourhood profiles. Paired tests revealed significant variations in the census-derived mortality rates between profile 1, associated with the lowest housing quality, and the three other profiles. These findings demonstrate the importance of housing quality as an important health risk factor. From a public health perspective, it should be a useful contribution for geographically targeted planning health policies, at the neighbourhood spatial level, which is the most appropriate administrative level for interventions.

Introduction

Rapid urbanization poses significant challenges in developing countries (Kessides, 2007). Urban planners have to deal with uncontrolled population growth leading to poverty, unemployment, promiscuity, spontaneous housing, and lack of necessary services (Korah et al., 2019) as well as subsequent health problems (Boadi et al., 2005). Well-being and health outcomes are generally better in urban than in rural areas (Yaya et al., 2019), but this urban advantage is mostly driven by the urban rich, and large disparities exist within the urban population (Dye, 2008; Günther and Harttgen, 2012; Gulyani et al., 2014). Socio-economic inequalities lead to increasing health inequalities in urban environments (Wagstaff, 2002; Quentin et al., 2014), especially in sub-Saharan Africa (SSA) where cities have the largest proportion of population living in informal and precarious settlements (Dos Santos et al., 2015; Günther and Harttgen, 2012). This situation can be observed in Dakar and its region (Senegal). For example, the urbanization of Pikine, a department of Dakar, has led to an uncontrolled urban sprawl, densely and heterogeneously populated areas, as well as unequal distribution of health care services (Salem, 1998). Borderon et al. (2014) characterized the habitat of the Dakar metropolitan area at a relatively fine scale (the census district level, N=1998) and revealed wide wealth and poverty inequalities. Varying risks of malaria infection within this area have also been associated with social vulnerability (Borderon and Oliveau, 2017).

In SSA cities, the lack of accurate and detailed health data is a major problem for the study of health as inequality-driven by rapid urbanization. Data on both health determinants and health outcomes are often lacking or are of poor quality (Quentin et al.,

2014; Satterthwaite *et al.*, 2019). Sample surveys, such as Demographic and Health Surveys, rarely make it possible to analyse variations in health indicators within cities, due to small sample sizes. Death registration systems are often incomplete, and only a few African cities, such as Antananarivo, Abidjan or Harare, have sufficiently developed civil registration systems providing robust mortality indicators (Dlodlo *et al.*, 2011; Masquelier *et al.*, 2019). In this context, national censuses fill an important gap, as they make it possible to study both mortality and the main socio-demographic characteristics of households. However, census data seldom include information on income, and when they do, the data are not sufficiently reliable and detailed to reveal intra-urban variations. Information on socio-professional categories is also problematic in contexts where a large portion of the population works in the informal sector. To overcome these limitations, proxies are used in order to analyse differences of socio-economic status and related health problems. One of the most straightforward proxies of socio-economic status is housing quality (Arias and De Vos, 1996; Bawah and Zuberi, 2004; Suglia *et al.*, 2011; Adjei and Kyei, 2013). This is a multidimensional concept, which results from a combination of components of built and natural environments, such as dwelling characteristics, household possessions and landscape attributes (Lanrewaju, 2012). More generally, adequate housing is a central component of productive, healthy, and meaningful lives, and a key social determinant of health and well-being (Tusting *et al.*, 2019).

In the literature, housing quality is evaluated through two main data sources: census or household surveys (Bawah and Zuberi, 2004; Suglia *et al.*, 2011; Adjei and Kyei, 2013), on the one hand, and remote sensing data (Thomson and Hardin, 2000; Rahman *et al.*, 2011) on the other. Questions regarding habitat characteristics are generally part of the census questionnaires. They provide a range of information such as type of house, number of rooms, building materials, occupancy status, level of comfort, access to safe water and sanitation. Using these variables to construct a proxy for economic status, Bawah and Zuberi (2004) created a composite poverty index that they employed in multivariate models to examine its association with childhood mortality in Botswana, Lesotho and Zambia. Remotely sensed data can be used as a complement to census data to map socio-economic indicators (Tapiador *et al.*, 2011; Gething *et al.*, 2015; Georganos *et al.*, 2019). Geographical information systems (GIS) technology and earth observation (EO) methods and techniques enable the use of satellite imagery to extract physical properties from the environment for urban studies (Bhatta *et al.*, 2010). This allows, for instance, mapping and monitoring urban sprawl and analysing associated environmental and economic impacts (Linard *et al.*, 2013). Very-high resolution (VHR) remotely sensed data allow for the delineation of objects such as buildings, roads, trees and are therefore particularly useful to derive detailed land cover and land use maps at the intra-urban level (Grippa *et al.*, 2017b; 2018a). Such maps also provide additional information related to housing quality through spatial metrics characterizing, for example the size, density and organization of buildings. Grippa and Georganos (2018) recently mapped the land cover and land use of the city of Dakar using a VHR satellite imagery. Based on an object-oriented approach to delineate different land cover classes (*e.g.*, vegetation, buildings, water) and spatial metrics, different land uses have been derived such as planned residential areas, deprived habitats or administrative, commercial and service areas. Extracting information related to housing quality from satellite imagery is much less expensive and less labour intensive than carrying out a census. It also has the advantage of producing spatially continuous data. However, it requires a high level of technical skills and the information derived from satellite imagery is associated with a higher level of uncertainty (Borderon *et al.*, 2014). Integration of census data and remotely sensed data has received attention, as both sources can be complementary with respect to mapping populations and settlements (Tatem *et al.*, 2007) and also provide information regarding the different components of housing quality (built and environmental aspects) (Borderon *et al.*, 2014).

In this study, we created fine scale housing quality maps for Dakar, and provide an original typology of neighbourhoods (N=1347). The innovative aspect of our approach lies in the combination of housing data from the last available census (2013) with VHR land use and land cover information developed by Grippa and Georganos (2018), which together cover various aspects of housing quality (*e.g.* dwelling characteristics, household possessions and landscape attributes). Our analysis is carried out for the four departments of the Dakar region and at the neighbourhood level, which is the most relevant level for health interventions. As an example of application, we also analyse the relation between housing quality, as defined by our housing quality components and profiles, and the crude mortality rate obtained from recent household deaths reported in the census. From a public health perspective, this study aims to help governments and other stakeholders engaged in urban planning to manage the risks caused by the high population growth and better understand spatial variations of mortality in the Dakar region in order to target health interventions correctly.

Materials and methods

Study area

The Dakar region is divided in four departments: Dakar, Pikine, Guediawaye and Rufisque (Figure 1). The department of Dakar, particularly the city centre, stands out from others thanks to its economic dynamism (UN-Habitat, 2008). Its large markets, industrial sector, companies and administrative buildings make it the driving force of Senegalese economy attracting migrants from other regions of the country. As the majority of SSA cities, strong demographic growth started after the independence of Senegal in 1960. Between the 1976 and 1988 censuses, its population increased from 0.9 to 1.5 million. In 2002, there were 2.2 million inhabitants and according to the National Agency for Statistics and Demography (ANSD), the last census in 2013 counted a population of 3.1 million (ANSD, 2013). This has resulted in high population densities and led to a demand for housing exceeding the existing supply. It also led to a marked unrestricted spread of the city into the area surrounding it and large-scale spontaneous settlements (Ndiaye, 2015). These settlements cover 42% of residential areas in Pikine, compared to 3% in Dakar (UN-Habitat, 2008). As a result, a large part of the population lives in poor housing conditions, especially in the periphery.

Data

Census data

Census data for the year 2013, recorded by enumeration area

(N=4189), were obtained from the Senegal census bureau of the ANSD. This information can be aggregated at different administrative levels, *i.e.* neighbourhoods (quartiers in French), communes, departments, region. We worked at the neighbourhood level (N=1347), the smallest administrative unit for which associated boundary data are available.

The census data provide the following sociodemographic information of each household member: age, sex, marital status, and household living conditions that includes a description of the inhabited house, goods and belongings and access to basic services, such as transport, sanitation and waste disposal. Heads of households are also asked about the number of deaths that occurred in the household in the 12 months preceding the census, *i.e.* from November-December 2012 to November-December 2013.

Remotely sensed data

Open access land cover (LC) and land use (LU) maps derived from Pléiades VHR satellite imagery, for the year 2015, were used in this study (Grippa *et al.*, 2018a). The spatial resolution is 0.5 m for LC map and street block level for LU map. Local segmentation and random forest classification according to Grippa *et al.* (2017b) allowed the identification of 11 LC classes (Figure 2A). The LC map was converted in a LU map using spatial metrics extracted at the street block level and divided into the following 8 LU classes:

planned residential, low-density planned residential, deprived residential, non-residential built-up, agricultural vegetation, natural vegetation, artificial ground surface and bare soil (Figure 2B). The LC and LU maps produced have an overall accuracy of 89.5% and 79%, respectively. These remote sensing data do not cover some parts of the eastern Dakar. Also, the Goree and Ngor islands are not shown in the LU map. The non-covered eastern zones are sparsely built-up areas in the urbanisation front.

Housing quality based on neighbourhood typology

Although there is a general consensus that variations in housing conditions reflect variations in living standards, there is no rule on the choice of variables that best identify housing conditions.

Here we combined data from the 2013 census and remote sensing (LC and LU) in order to best characterize housing quality. More precisely, we used the framework developed in Bawah and Zuberi (2004), in combination with Dakar's detailed LC and LU data (Grippa *et al.*, 2018). Table 1 summarizes the 16 variables and associated data sources used by neighbourhood in the subsequent typology. A subset of 8 variables was derived from the census housing database, based on multicollinearity tests. Among the 8 variables used, 6 were individual (household ownership, type of roof (concrete slab, tiled roof), type of walls (improved cement), type of sewage system and availability of a private tap), while 2 related to the type of floor covering high-quality and poor quality were combined.
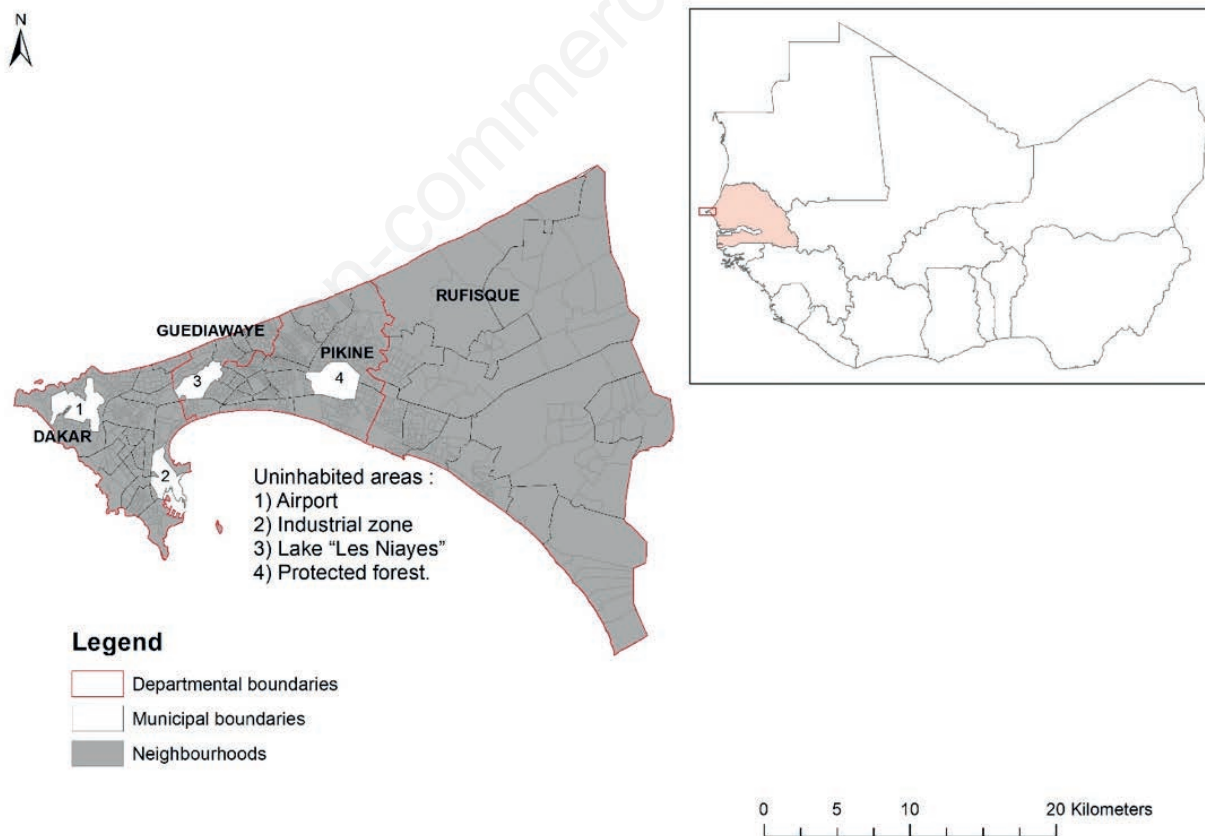

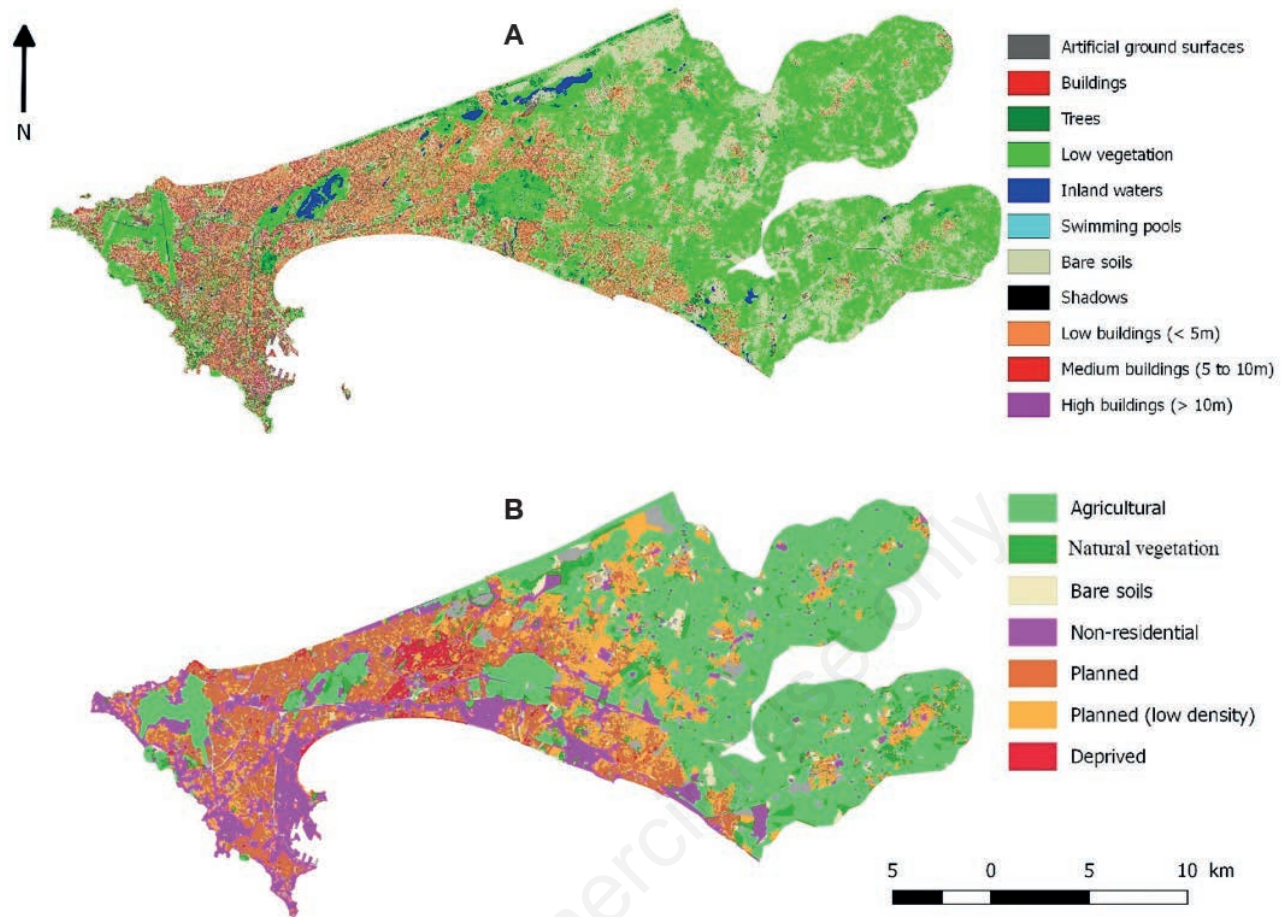**Figure 1. Geographical location of Dakar and its administrative units.**

Figure 2. Land cover (A), and land use (B) maps of Dakar (data source: https://zenodo.org/record/1290800#.W7xWQ_a2w2w).

Table 1. Variables extracted from census and satellite-derived maps and used in the typology of neighbourhoods.

Variable	Description	Source
Owner	Proportion of household heads being owners	Census
Concrete slab	Proportion of households with roofs made of concrete slabs	Census
Tiled roof	Proportion of households with tiled roofs	Census
Improved cement	Proportion of households walls made of cement with tiles or cement and marble	Census
High-quality floor covering	Proportion of households with floors covered with tiles or carpet	Census
Poor-quality floor covering	Proportion of households with floors covered with sand or clay	Census
Sewage system	Proportion of households connected to the sewage network	Census
Private tap	Proportion of households relying on water from a house tap or mineral water	Census
Planned habitat	Proportion of planned residential LU	LU map
Deprived habitat	Proportion of unplanned residential LU	LU map
Non-residential built-up	Proportion of LU without residences used for administrative and commercial services	LU map
Agricultural vegetation	Proportion of agricultural areas	LU map
Natural vegetation	Proportion of natural vegetation	LU map
Bare soil	Proportion of bare soil	LC map
Population density	Population per total area	Census
Built-up density	Population per built-up area	LC map, census

LU, land use; LC, land cover.

They were selected because: i) the coefficient of variation was lower than individual variables; and ii) Pearson correlation coefficient between categorical variables characterizing the type of floor was high, especially those related to the quality of covering. LC raster map was reclassified in order to extract several LC classes. Zonal statistics and tabulate area tools were used to compute the proportion of built-up area, planned residential, deprived residential, non-residential, bare soil, agricultural vegetation and natural vegetation by neighbourhood. In addition, two different population density metrics were extracted for each neighbourhood using population counts from the census data: i) the total population density, where population counts are divided by the total neighbourhood area; and ii) the built-up population density, where population counts are divided by the built-up area derived from the LC map (Salem, 1998).

A principal component analysis (PCA) was used to reduce the dimensionality of the dataset and create housing quality indices. PCA has been widely employed in social sciences for creating indices for issues related to poverty and health inequalities (Lo and Faber, 1997; Bawah and Zuberi, 2004; Li and Weng, 2007). Using the FactoMineR package of the R statistical software (Lê *et al*., 2008), PCA was applied to the 16 variables listed in Table 1. After extracting the factor loadings, PCA showed the correlation coefficients between each component and each variable, which allowed us to score neighbourhoods on each component. This kind of score also refer to their level of contribution in the formation of this component. Housing quality indices were derived from the neighbourhood coordinates in the multi-dimensional space. We mapped PCA results and analysed spatial distribution of neighbourhood characteristics.

A k-means hierarchical classification was carried out based on the PCA housing quality indices in order to create a typology of neighbourhoods (N=1347). The k-means is an algorithm for partitioning the hierarchical tree, which allows a more optimized aggregation of neighbourhoods, thus creating homogeneous neighbourhood clusters. K-means use an iterative process to group observations according to similarities in the mean values of the factor scores (Owens, 2012). Test values (V-tests) were used to describe the clusters. The V-test statistic measures the deviation of the cluster mean from the overall mean in number of standard deviations. Higher absolute values indicate a stronger importance of the component to characterize the cluster.

Relationships between housing quality and mortality rates

To assess whether mortality is associated with housing quality, we analysed the distribution of overall mortality rates through both neighbourhood components and the profiles obtained from the hierarchical clustering. Statistical tests, including the Kruskal-Wallis test and paired tests were used to test for mortality rate differences between clusters. Pearson's correlation coefficients were computed for comparisons between crude mortality rate and PCA's derived housing quality components.

Results

Housing quality based on neighbourhood typology

The four first components were extracted from the PCA results. Together these four components captured 59% of the total variance, with the first explaining 27% of the total variation, the second 13%, the third 10% and the fourth 9%. The four components had an eigenvalue >1. For further information, we examined the correlation coefficients between the variables in the original data matrix and the components (Table 2). To deduce that a variable is associated with a principal component, Comrey and Lee (1992) demand that the correlation coefficients should be higher than 0.45 or lower than –0.45, while coefficients >0.71 or <0.71 are considered excellent, 0.63 very good and 0.45 fair. For this reason, only correlation coefficients above 0.45 or below –0.45 are shown in the table.

Component 1 showed strong positive loadings on the variables '*concrete slab*', '*high quality covering*', '*private tap*' and '*planned habitat*' but was negatively correlated to the variables '*tiled roof*' and '*deprived habitat*'. Higher scores on this component point to households living in comfortable houses, where the roof is made of solid materials, having water on tap and well equipped with other housing facilities such as carpeted floors. We labelled this component as '*high-quality housing*'. The '*planned habitat*' variable refers to the rectangular form of the housing blocks, which confirms the presence of straight-lined streets that is suitable for car traffic and service provision such as household waste collection.

Table 2. Correlation coefficients between input variables and four components extracted from principal component analysis.

Component	1	2	3	4
Owner			0.55	
Concrete slab	0.89			
Tiled roof	–0.88			
Improved cement				
High-quality floor covering	0.86			
Poor-quality floor covering			0.66	
Sewage system			–0.72	
Private tap	0.65			
Planned habitat	0.53	–0.60		–0.50
Deprived habitat	–0.70			
Non-residential built-up		0.47		0.74
Agricultural vegetation		0.65		
Natural vegetation		0.49		
Bare soil				
Population density		–0.69		
Built-up density				
Eigenvalue	4.05	1.93	1.48	1.37
Percentage of variance	27.04	12.91	9.90	9.13
Cumulative percentage of variance	27.04	39.95	49.86	59

Variables with noteworthy positive loadings on Component 2 were '*agricultural vegetation*', '*natural vegetation*' and '*non-residential built-up*' areas. Component 2 was also negatively correlated with '*planned habitat*' and '*population density*'. Therefore, this component refers to households living in less densely populated neighbourhoods, with a high proportion of both natural and agricultural vegetation and to a lesser extent a higher proportion of non-residential built-up areas. Component 2 was not correlated with any variable related to household equipment and building materials. We used the term '*vegetated environment*' when referring to this component. For Component 3, high positive loadings related to the variables '*poor quality covering*' and '*owner*'. It was also negatively correlated with the variable '*sewage system*'. This component identifies households that lack comfort facilities, where sanitation services do not exist, but those who live there own their homes. Even if the interpretation of this component is less intuitive, such households usually correspond to more traditional housing and we labelled this component as '*poor-quality properties*'. Component 4 showed a strong positive loading on the variable '*non-residential built-up*', indicating neighbourhoods dominated by administrative, commercial and service activities. This component is also negatively correlated to the variable '*planned habitat*' and was labelled as '*business areas*'.

Figure 3 shows maps of the four defined components: high-

quality housing, vegetated environments, poor-quality properties and business areas. In general, western neighbourhoods present higher-quality housing and are clearly opposed to the eastern neighbourhoods (Figure 3A). However, fine-scale spatial variations can also be observed, which mitigate this overall duality. Although the contribution of central Dakar in shaping the first dimension is prominent, high-quality housing is also seen in the periphery, notably in the Departments of Guediawaye and Rufisque. Vegetated neighbourhoods are mainly located on the eastern urbanization front (Figure 3B). In the centre, Component 2 also highlights pockets of vegetation around the protected forest and the wetlands, such as the lake '*Les Niayes*' (Figure 1) and along the western coast. Component 3 represents the poor-quality properties and higher scores are located in the rural areas of Rufisque (Figure 3C). Its spatial distribution in the Southwest also reveals fine-scale nuances, with a striking opposition between positive and negative coordinates among some municipalities of the department of Dakar (*e.g.*, between the Plateau and Medina neighbourhoods). Figure 3D shows that administrative, commercial and services are mainly concentrated in the city centre (Dakar department) and along the southern coast. The east-west opposition is clear for Component 4, with tertiary activities being poorly represented in the outlying eastern neighbourhoods.
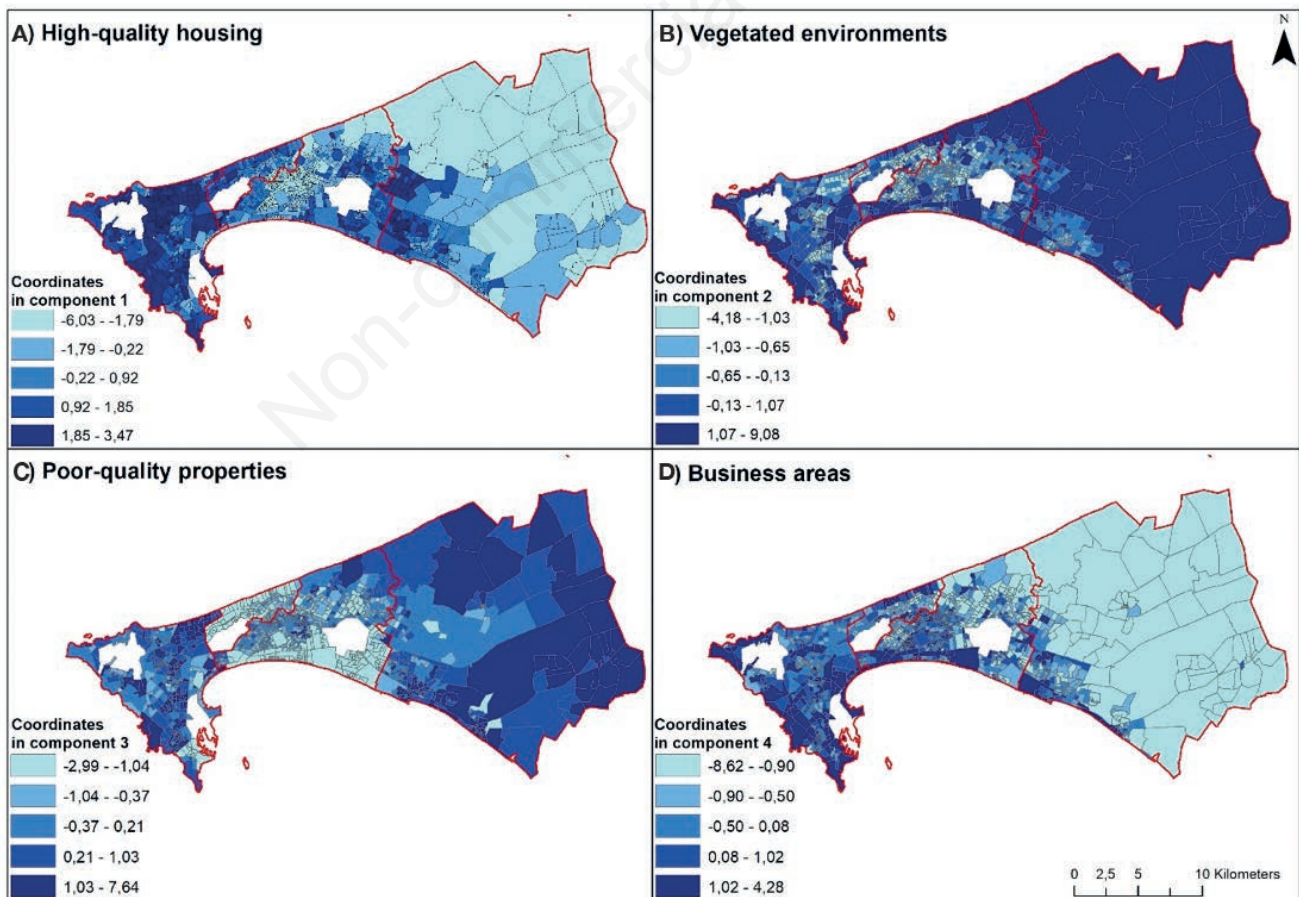

Figure 3. A-D) Spatial distribution of the four components defined by the principal component analysis.

Turning to hierarchical classification to further distinguish housing characteristics and uncover distinct spatial patterns, four clusters were selected based on visual interpretation of the dendrogram. Table 3 presents V-test results and Table 4 shows the significance levels of V-test statistics for the 16 variables used in the PCA analysis. Cluster 1 included neighbourhoods characterized by a negative association with Component 1, *i.e.* lower-quality housing neighbourhoods. Here a large proportion of the population lives in houses built with tiles in deprived residential areas. This cluster accounted for almost one quarter of total neighbourhoods (306 out of 1325) and was mainly located in Pikine (241 out of 306) (Figure 4).

Table 3. v-test statistic of components by cluster.

PCA Cluster component	1	2	3	4
High-quality housing	−26.77	−8.58	9	20.59
Vegetated environments	−9.96	24.06	14.67	−16.37
Poor-quality properties	-	4.75	−4.74	-
Business areas	5.82	−15.01	23	−13.93

Table 4. Cluster significance level.

Variable Cluster	1	2	3	4
Owner	--	+	-	+
Concrete slab	--	--	++	+
Tiled roof	+++	++	--	--
High quality covering	--	--	+	+
Poor quality covering	-	--	+	
Sewage system	+	+	-	-
Private tap	--	--	+	+
Planned habitat	-	--	-	+++
Deprived habitat	+++	--	--	--
Non-residential built-up	--	--	+++	--
Agricultural vegetation	--	+++	--	--
Natural vegetation	--	+++	-	--
Bare soil	---	+++	+	-
Population density	++	---	--	+

The table is based on v-test computing variable's percentage deviation of the mean: +, moderate positive association (0-45%); ++, high positive association (45-80%); +++, very high positive association (80-100%); The negative results are symbolized by - using the same levels of strengths.
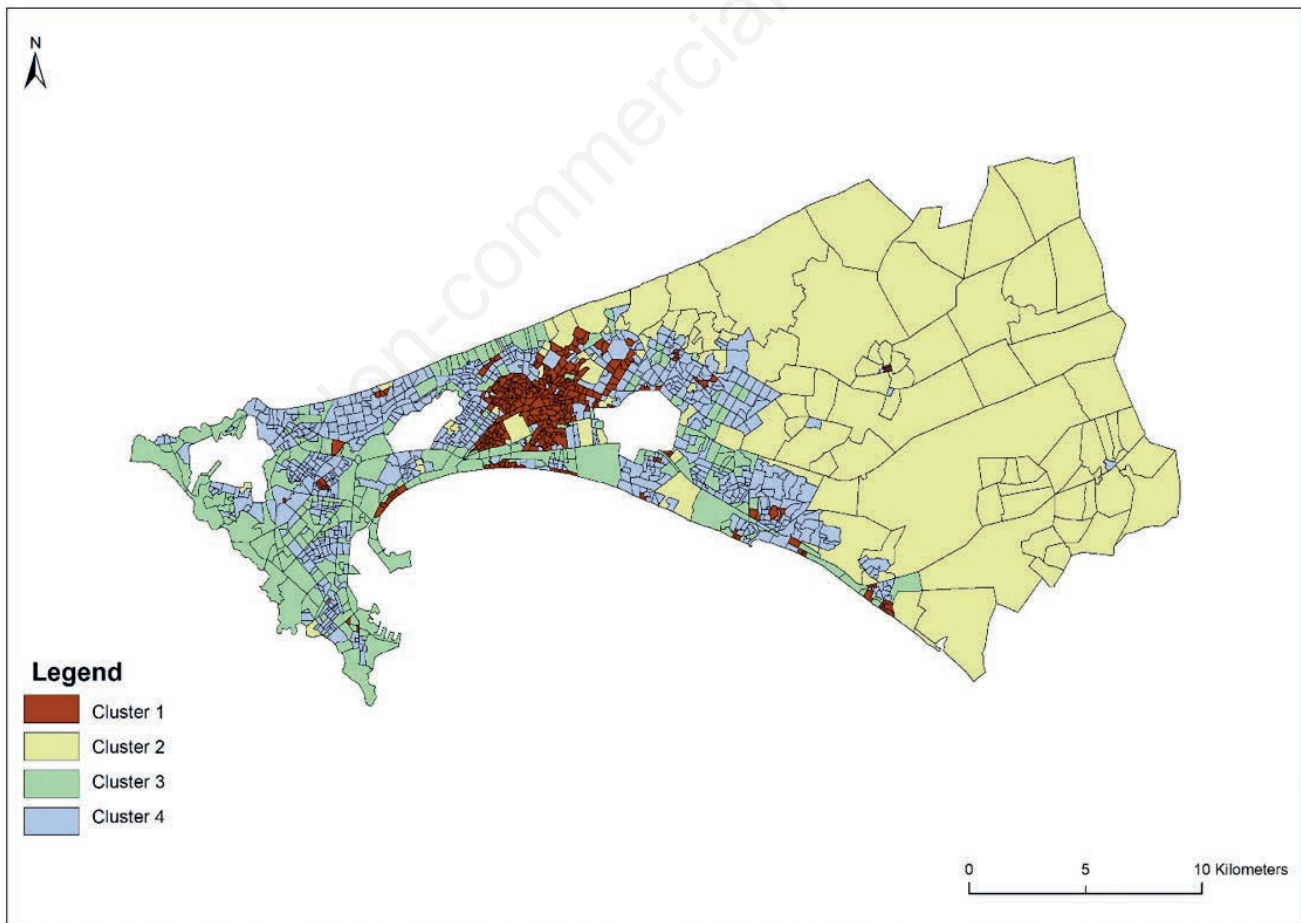

Figure 4. Distribution of neighbourhoods by clusters.

Cluster 2 was positively associated with PCA-component 2, *i.e.* vegetated environments. It had the largest proportion of home-owners (40% of households) and the fewest number of neighbourhoods (116 out of 1325), mainly located in the eastern and less densely populated areas. Table 4 also highlights the importance of bare soil in addition to natural and agricultural vegetation in Cluster 2.

Cluster 3 mainly refers to PCA-component 4 (business areas), but combines both non-residential and residential functions. Half of the land use areas here are for administrative, commercial and service purposes. As for the household dwellers, 76.2% have private tap and 73.5% have high-quality floor coverage on their premises. This cluster comprises 233 neighbourhoods.

Cluster 4 included neighbourhoods characterized by a positive association with Component 1, *i.e.* higher-quality housing. This class is characterized by a high proportion of residential planned habitats with 80% of households having concrete roof type, 69.3% having high-quality soil coverage and 74.7% using safe water for drinking. This cluster had the largest number of neighbourhoods (670 out of 1325). Clusters 3 and 4 spread outside the city centre, in the department of Rufisque.

Relationships between housing quality and mortality rates

We examined the distribution of crude mortality rate in the four neighbourhood profiles defined above. The highest mortality rate was reported in Cluster 1 (Figure 5). The Kruskal-Wallis test showed significant differences between clusters (P<0.001).

However, paired tests revealed that differences in averages were only significant between Cluster 1 (with mean 5.82 and 95% confidence interval 5.47-6.16) and the other clusters (mean 4.27 and 95% confidence interval 3.76-4.79).

Differences in mortality rates between Clusters 2, 3 and 4 were not significant. Pearson's correlation between mortality and PCA's derived components showed a highly significant negative correlation between crude mortality rate and Component 1 (P<0.001). In other words, high housing-quality neighbourhoods reported lower mortality rates. Component 2, which depicts vegetated environments, was also significantly negatively correlated with mortality (P=0.005). The relationships between the crude death rate and the remaining components were not statistically significant (P=0.79 for the correlation between the mortality rate and Component 3; P=0.13 for the correlation between the mortality rate and Component 4).

Discussion

The four clusters identified in this study are based on a combination of census and remotely sensed variables and describe both the built and environmental components of housing quality. This neighbourhood typology provides a picture of housing quality as complete as possible, regarding the accuracy of LCLU maps, the quality of census data and the correlation found with the crude death rate.
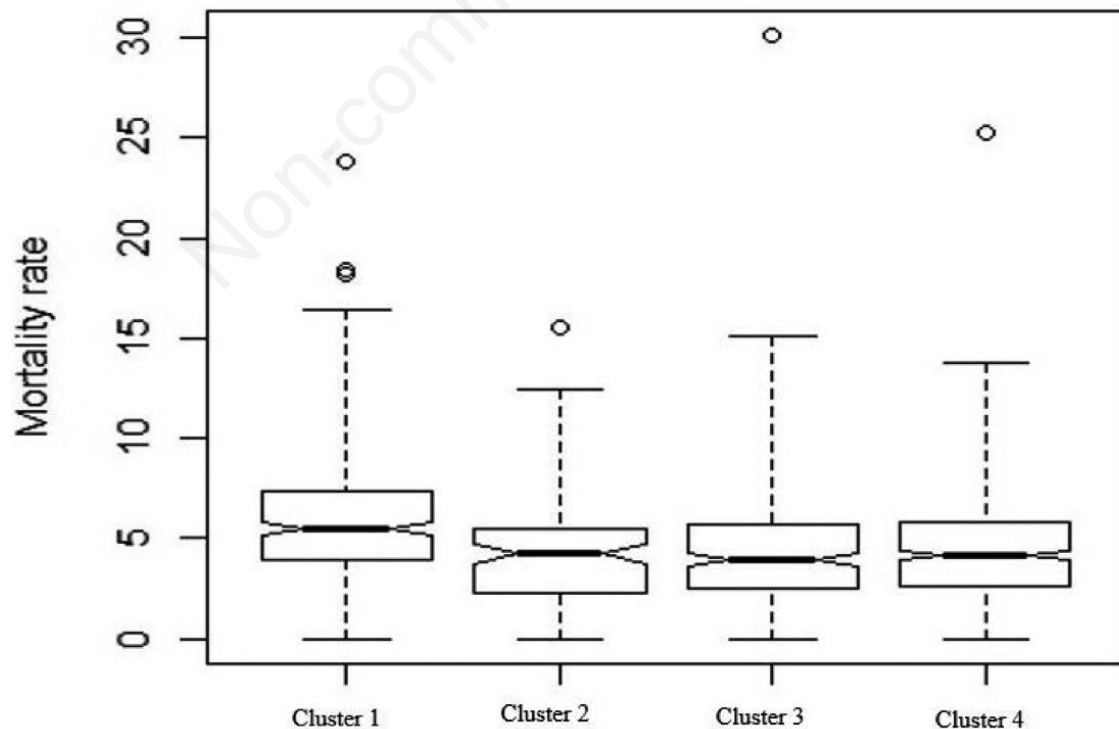

Figure 5. Distribution of crude mortality rates by cluster.

The main objective of this study was to develop a typology of neighbourhoods based on housing quality in Dakar and it appears that population density plays an important role in the resulting typology. This information allowed us to distinguish between two profiles of peri-urban neighbourhoods: the small dense spaces in the suburb of Pikine, as already observed by Salem (1998), and the neighbourhoods on the front line of urbanization characterized by a lower population density.

Our results confirm the previous typology carried out by Borderon (2014) based on the 2002 census data, which highlighted the fine-scale heterogeneities of housing quality across the city. However, as shown by our study, noteworthy developments have occurred since the 2002 census. The residential housing profile (cluster 4) expanded from the city centre to the periphery, probably as a result of urban integration strategies that allowed access to ownership for a large part of the population.

The typology constructed in this study can be cross-tabulated with other variables, such as subjective poverty and access to health care, after controlling for other socio-demographic variables, such as education level. The level of education is recognized as an important factor explaining socio-economic and health differences (Kofie et al., 2008; Adjei and Kyei, 2013), and closely related to housing quality (Tusting et al., 2019). The comparison between housing quality indices and crude mortality rates provided expected results. The variation in the crude mortality rate was shown to be significantly different between neighbourhood profiles, and high mortality is associated with low-quality housing (Cluster 1). However, the crude mortality rate suffers from certain deficiencies that can lead to misinterpretations. As it does not take into account differences in the age structure of the population, it can therefore mask differences, for example between child and adult mortality. Crude mortality is also sensitive to variations in the age structure of the population (ANSD, 2013), so age-specific mortality rates are preferable to the crude mortality rate to better reflect health inequalities. However, the excess zeros problem is raised when analysing age-specific mortality rates at disaggregated levels.. In addition to age, information on the causes of death is also essential for the analysis of health inequalities, as different neighbourhoods may be affected differently by infectious or chronic diseases. Another limitation generally affecting mortality data is the risk of underreporting deaths. Ideally, mortality statistics from the census should be triangulated by estimates from the death registration system. The Senegalese system should be strengthened to increase the coverage and the precision of spatial information on places of residence of the deceased.

Our study adds to the existing literature, showing evidence of association between housing quality and health, including child mortality, infectious and chronic diseases such as malaria, respiratory illness, obesity, parasitic disease. Jankowska et al. (2013) in Accra, Ghana showed that child mortality is connected to environmental factors such as housing quality, slum-like conditions and the presence of vegetation in the neighbourhood. Yaya et al. (2018) examined in Nigeria the association between household characteristics, such as toilet facilities, water quality, access to electricity, existence and quality of walls, roofs and floor in Nigerian dwellings on the one hand, and diarrhoea amongst various under-five children on the other. The authors concluded that substandard living conditions can contribute to the increasing burden of diarrhoea. Despite using different measures of housing quality and various health data, the studies published in this focused literature all support the idea that poor housing quality is associated with poor health. Some authors highlight that housing improvements can lead to health improvements (Vaid and Evans, 2017) and that improved housing quality can be a key component of health policy, notably in Africa (Herrin et al., 2013).

In our analysis of the links between spatial variations in habitat quality and variations in mortality rates, it is essential to consider the risks of ecological errors and the modifiable areal unit problem (MAUP) (Ayubi and Safiri, 2018). These problems may arise here because individual census data have been aggregated to a coarser geographical level (the neighbourhood). Statistically, while correlations are easy to establish, causality is less obvious, as correlations at the neighbourhood level - whatever the spatial detail - may differ from those at the individual level. Results and interpretations may change as the scale of spatial representation changes. Also, combining census data from 2013 with satellite data from 2015 presents a slight risk of ecological errors due to the rapid evolution of urban land use. Nevertheless, the higher crude mortality rate in Cluster 1 as well as lower mortality in higher housing quality provides important insights in the relationship between urban areas and health in Dakar.

Conclusions

Our study demonstrates that combining census and remote sensing data is feasible and can improve our understanding of urban patterns. The use of VHR satellite imagery provided a more precise description of the characteristics of the urban landscape in Dakar, which complemented the more detailed thematic information provided by census data. Together, these two data sources allowed a better identification of housing quality variations that need to be considered to guide public health interventions.

References

Adjei PO-W, Kyei PO, 2013. Linkages between income, housing quality and disease occurrence in rural Ghana. J Hous Built Environ 28:35-49.

ANSD (Agence Nationale de la Statistique et de la Démographie), 2013. Rapport définitif RGPHAE-2013, 418 pp. Available from: http://www.ansd.sn/ressources/RGPHAE-2013/resultats-definitifs.htm

Arias E, De Vos S, 1996. Using housing items to indicate socioeconomic status: Latin America. Soc Indic Res 38:53-80.

Ayubi E, Safiri S, 2018. Geographic variation in morbidity and mortality of cerebrovascular diseases in Korea during 2011-2015: bias due to spatial autocorrelation and modifiable areal unit problem. J Stroke Cerebrovasc Dis Off J Natl Stroke Assoc 27:1123.

Bawah AA, Zuberi Z, 2004. Socioeconomic status and child mortality: an illustration using housing and household characteristics from African census data. Afr Popul Stud 19:9-29.

Bhatta B, Saraswati S, Bandyopadhyay D, 2010. Urban sprawl measurement from remote sensing data. Appl Geogr

30:731-40.

Boadi K, Kuitunen M, Raheem K, Hanninen K, 2005. Urbanisation without development: environmental and health implications in African cities. Environ Dev Sustain 7:465-500.

Borderon M, Oliveau S, 2017. Vulnérabilités sociales et changement d'échelle. Espace Popul Sociétés Space Popul Soc 2016/3. Available from: http://journals.openedition.org/eps/7012

Borderon M, Oliveau S, Machault V, Vignolles C, Lacaux J-P, N'Donky A, 2014. Qualifier les espaces urbains à Dakar, Sénégal. Cybergeo Eur J Geogr. Available from: http://journals.openedition.org/cybergeo/26250

Comrey AL, Lee HB, 2016. A first course in factor analysis, 2nd edn. Psychology Press & Routledge (Taylor & Francis), London, UK.

Dlodlo RA, Fujiwara PI, Hwalima ZE, Mungofa S, Harries AD, 2011. Adult mortality in the cities of Bulawayo and Harare, Zimbabwe: 1979-2008. J Int AIDS Soc 14:S2-S2.

Dos Santos S, Rautu I, Diop M, Abdou Illou MM, Ndonky A, Le Hesran J-Y, Lalou R, 2015. The influence of environmental factors on childhood fever during the rainy season in an African city: a multilevel approach in Dakar, Senegal. Popul Environ 36:429-51.

Dye C, 2008. Health and urban living. Science 319:766-9.

Georganos S, Gadiaga AN, Linard C, Grippa T, Vanhuysse S, Mboga N, Wolff E, Dujardin S, Lennert M, 2019. Modelling the wealth index of demographic and health surveys within cities using very high-resolution remotely sensed information. Remote Sens 11:2543.

Gething PW, Tatem AJ, Bird T, Burgert C, 2015. Creating spatial interpolation surfaces with DHS data. DHS Spatial Analysis Reports No. 11. Rockville, Maryland, USA: ICF International. Available from: https://dhsprogram.com/pubs/pdf/SAR11/SAR11.pdf

Grippa T, Georganos S, 2018. Dakar land use map at street block level (Zenodo). Available from: https://zenodo.org/record/1291389#.YIfSILUzY2w

Grippa T, Lennert M, Beaumont B, Vanhuysse S, Stephenne N, Wolff E, 2017a. An open-source semi-automated processing chain for urban object-based classification. Remote Sens 9:358.

Grippa T, Georganos S, Vanhuysse SG, Lennert M, Wolff E, 2017b. A local segmentation parameter optimization approach for mapping heterogeneous urban environments using VHR imagery. In: W. Heldens, N. Chrysoulakis, T. Erbertseder, and Y. Zhang (Eds.), Remote sensing technologies and applications in urban environments II. SPIE. Warsaw, Poland, p. 20.

Grippa T, Georganos S, Zarougui S, Bognounou P, Diboulo E, Forget Y, Lennert M, Vanhuysse S, Mboga N, Wolff E, 2018. Mapping urban land use at street block level using open street map, remote sensing data, and spatial metrics. ISPRS Int J Geo-Inf 7:246.

Gulyani S, Bassett EM, Talukdar D, 2014. A tale of two cities: a multi-dimensional portrait of poverty and living conditions in the slums of Dakar and Nairobi. Habitat Int 43:98-107.

Günther I, Harttgen K, 2012. Deadly cities? spatial inequalities in mortality in sub-Saharan Africa. Popul Dev Rev 38:469-86.

Herrin WE, Amaral MM, Balihuta AM, 2013. The relationships between housing quality and occupant health in Uganda. Soc Sci Med 81:115-22.

Jankowska MM, Benza M, Weeks JR, 2013. Estimating spatial inequalities of urban child mortality. Demogr Res 28:33-62.

Kessides C, 2007. The urban transition in Sub-Saharan Africa: challenges and opportunities. Environ Plan C Gov Policy 25:466-85.

Kofie RY, Attua EM, Nabila JS, 2008. Poverty and socio-economic consequences of Buruli ulcer (Mycobacterium ulcerans) in the Ga West District of Ghana. Nor Geogr Tidsskr - Nor J Geogr 62:210-21.

Korah PI, Matthews T, Tomerini D. 2019. Characterising spatial and temporal patterns of urban evolution in Sub-Saharan Africa: the case of Accra, Ghana. Land Use Policy 87:104049.

Lanrewaju F, 2012. Urbanization, housing quality and environmental degeneration in Nigeria. J Geogr Reg Plan 5:422-9.

Lê S, Josse J, Husson F, 2008. FactoMineR: an R package for multivariate analysis. J Stat Softw 25.

Li G, Weng Q, 2007. Measuring the quality of life in city of Indianapolis by integration of remote sensing and census data. Int J Remote Sens 28:249-67.

Linard C, Tatem AJ, Gilbert M, 2013. Modelling spatial patterns of urban growth in Africa. Appl Geogr 44:23-32.

Lo CP, Faber BJ, 1997. Integration of landsat thematic mapper and census data for quality of life assessment. Remote Sens Environ 62:143-57.

Masquelier B, Pison G, Rakotonirina J, Rasoanomenjanahary A, 2019. Estimating cause-specific mortality in Madagascar: an evaluation of death notification data from the capital city. Popul Health Metr 17:8.

Ndiaye I, 2015. Étalement urbain et différenciation sociospatiale à Dakar (Sénégal). Cah Géographie Qué 59:47-69.

Owens A, 2012. Neighborhoods on the rise: a typology of neighborhoods experiencing socioeconomic ascent. City Community 11:345-69.

Quentin W, Abosede O, Aka J, Akweongo P, Dinard K, Ezeh A, Hamed R, Kayembe PK, Mitike G, Mtei G, 2014. Inequalities in child mortality in ten major African cities. BMC Med 12:95.

Rahman A, Kumar Y, Fazal S, Bhaskaran S, 2011. Urbanization and quality of urban environment using remote sensing and GIS techniques in East Delhi-India. J Geogr Inf Syst 03:62.

Salem G, 1998. La santé dans la ville: géographie d'un petit espace dense. Karthala Editions, Pikine, Sénégal.

Satterthwaite D, Sverdlik A, Brown D, 2019. Revealing and responding to multiple health risks in informal settlements in Sub-Saharan African cities. J Urban Health 96:112-22.

Suglia SF, Duarte CS, Sandel MT, 2011. Housing quality, housing instability, and maternal mental health. J Urban Health 88:1105-116.

Tapiador FJ, Avelar S, Tavares-Corrêa C, Zah R, 2011. Deriving fine-scale socioeconomic information of urban areas using very high-resolution satellite imagery. Int J Remote Sens 32:6437-56.

Tatem AJ, Noor AM, Hagen C, von Gregorio AD, Hay SI, 2007. High resolution population maps for low income nations: combining land cover and census in East Africa. PLoS One 2:e1298.

Thomson CN, Hardin P, 2000. Remote sensing/GIS integration to identify potential low-income housing sites. Cities 17:97-109.

Tusting LS, Bisanzio D, Alabaster G, Cameron E, Cibulskis R, Davies M, Flaxman S, Gibson HS, Knudsen J, Mbogo C, Okumu FO, von Seidlein L, Weiss DJ, Lindsay SW, Gething PW, Bhatt S, 2019. Mapping changes in housing in sub-Saharan Africa from 2000 to 2015. Nature 568:391-4.

UN-Habitat, 2008. Sénégal: Profil urbain de Dakar; 36 pp. Available from: https://unhabitat.org/books/senegal-profil-urbain-de-dakar-french/

Vaid U, Evans GW, 2017. Housing quality and health: an evaluation of slum rehabilitation in India. Environ Behav 49:771-90.

Wagstaff A, 2002. Poverty and health sector inequalities. Bull World Health Organ 80:97-105.

Yaya S, Hudani A, Udenigwe O, Shah V, Ekholuenetale M, Bishwajit G, 2018. Improving water, sanitation and hygiene practices, and housing quality to prevent diarrhea among under-five children in Nigeria. Trop Med Infect Dis 3:41.

Yaya S, Uthman OA, Okonofua F, Bishwajit G, 2019. Decomposing the rural-urban gap in the factors of under-five mortality in sub-Saharan Africa? Evidence from 35 countries. BMC Public Health 19:616.