University of Namur

DOCTOR OF SCIENCES

Spatial analysis of health inequalities in Dakar, Senegal

GADIAGA, Assane

Award date:
2022

Awarding institution:
University of Namur

SPATIAL ANALYSIS OF HEALTH INEQUALITIES IN DAKAR, SENEGAL

Presented by :

Assane Niang Gadiaga

Supervisor:

Prof. Catherine Linard

Co-supervisor:

Prof. Mouhamadou Diallo

January 2022

Thesis jury:

Sabine Henry (Présidente – Université de Namur)

Catherine Linard (Promotrice – Université de Namur)

Mouhamadou Diallo (Co-promoteur – Université Cheikh Anta Diop)

Bruno Masquelier (Université Catholique de Louvain)

Gérard Salem (Université Paris Ouest Nanterre La Défense)

Niko Speybroeck (Université Catholique de Louvain)

List of abbreviations

ACS	Administrative, Commercial and Service
ANSD	Agence Nationale de la Statistique et de la Démographie (National Agency for Statistics and Demography)
AIDS	Acquired Immunodeficiency Syndrome
ARES-CCD	Académie de Recherche et d'Enseignement Supérieur – Commission de la Coopération au Développement
AUC	Area Under the Curve
BRT	Boosted Regression Trees
COVID-19	Coronavirus Disease 2019
CV	Cross validation
GLM	Generalised Linear Model
HIV	Human Immunodeficiency Virus
DHS	Demographic and Health Surveys
DHIS 2	Demographic and Health Information Software 2
GHE	Global Health Estimates
GIS	Geographical Information System
GWR	Geographically Weighted Regression
GWPR	Geographically Weighted Poisson Regression
IRR	Incidence Rate Ratios
LC/LU	Land Cover/Land Use
MAUP	Modifiable Areal Unit Problem
MCRC	Main Civil Registration Center
UN	United Nations
PCA	Principal Component Analysis
RGPHAE	Recensement Général de la Population et de l'Habitat, de l'Agriculture et de l'Elevage
RMSE	Root Mean Square Error
SARS-Cov-2	Severe Acute Respiratory Syndrome coronavirus 2
SSA	Sub-Saharan Africa
VHR	Very-High Resolution

WHO	World Health Organisation
WMR	World Malaria Report

Préambule

Cette thèse de doctorat s'inscrit dans le cadre du projet ASSESS (Améliorer les statistiques de décès pour mieux suivre les évolutions sanitaires dans la région de Dakar, http://assess-sn.org/), financé par l'Académie de Recherche et d'Enseignement Supérieur de la fédération Wallonie-Bruxelles (ARES-CDD, https://www.ares-ac.be). L'objectif global du projet ASSESS est de contribuer à la baisse de la mortalité dans l'agglomération Dakaroise et à la réduction des inégalités de santé. Son objectif spécifique est de mettre à la disposition des autorités sénégalaises des informations sanitaires plus complètes, désagrégées et contextualisés sur les décès et leurs principales causes. Les institutions impliquées dans ce projet de recherche sont : l'Université Catholique de Louvain (coordonnateur Nord), l'Université de Namur (partenaire Nord), l'Agence Nationale de la Statistique et de la Démographie (coordonnateur Sud) et l'Université Cheikh Anta Diop (partenaire Sud).

Les résultats attendus de ce projet de recherche pour le développement sont les suivants : (R1) des recommandations sont émises pour améliorer la couverture de l'enregistrement des décès et l'exploitation des données d'état civil à des fins statistiques, (R2) la faisabilité et les avantages d'un relevé des causes de décès au sein des centres d'état civil sont démontrés, (R3) les informations sanitaires et démographiques sont mieux mises en relation avec les données de contexte. La présente recherche s'inscrit dans ce dernier point. Dans l'atteinte de ce résultat de recherche, d'autres synergies ont pu être développées. Elles concernent le projet MAUPP (http://maupp.ulb.ac.be) (SR/00/304) et le projet REACT (http://react.ulb.be/) (SR/00/337), tous deux financés par le programme STEREO-III de la Politique Scientifique Fédérale belge (BELSPO), ainsi que la collaboration avec le laboratoire Dynamiques Territoriales et Santé du département de Géographie de l'Université Cheikh Anta Diop.

Summary

Sub-Saharan African cities are highly heterogeneous, which makes the analysis of spatial health inequalities particularly interesting and challenging. Indeed, cities offer both the best and the worst environments for health and well-being. Multiple determinants converge to influence the health status of city dwellers, and positive and negative influences tend to cluster according to the specific neighbourhood or place within the city. The objective of this doctoral research is to analyse the spatial health inequalities in Dakar. In a health geography perspective, a systemic approach is developed to understand these spatial inequalities. Based on this approach, states of health are conceived as the result of the interplay between population, habitat and behaviour.

Available data sources were first examined to assess their potential for spatial analyses. The usefulness of civil registration and DHIS 2 data for spatial analyses is hampered by several deficiencies. Such deficiencies are overcome by census data, which can provide spatially representative health indicators that cover the whole region of Dakar. Efforts were made to ingrate the 2013's census data into a GIS system in order to conduct spatial analyses. In addition to census data, survey data and Land Cover/Land Use data derived from very-high resolution satellite imagery were used.

A typology of neighbourhoods was carried out and revealed the extreme heterogeneity of Dakar in terms of living conditions. These heterogeneities are reflected in health conditions, with higher crude mortality rates in spontaneous settlements, and lower crude mortality rates in residential neighbourhoods. A deeper understanding of these inequalities is gained through the decomposition of the crude mortality rate into age-specific mortality rates. Spatial autocorrelation analyses revealed presence of clusters in the spatial distribution of mortality for both child-adolescent, adult and the elderly age groups. A geographically weighted regression was used to model the impact of contextual risk factors in spatial variations of age-specific mortality. Determinants of mortality vary both across age-groups and across space, except for population density that is consistently positively associated with age-specific mortality rates.

The spatial distribution of COVID-19 infections is also clustered, with higher incidence rates in both western and eastern neighbourhoods. Different models highlighted population density as the most influential variable in explaining spatial variations in COVID-19 infection, followed by the highly connected administrative, commercial and service areas. The predicted spatial distribution of COVID-19 infections revealed a higher probability of infection in the western parts of Dakar, especially in areas characterized by a high population density and a high connectivity.

The influence of contextual risk factors on health operates through many indirect and complex mechanisms, and thus caution is needed when drawing conclusions about mortality or morbidity determinants. Above all, these results are useful indicators of mortality differentials and differentials in their determinants. From a public health perspective, results can help to develop geographically targeted interventions.

Résumé

L'hétérogénéité spatiale des villes rend particulièrement intéressante l'étude des inégalités spatiales de santé en milieu urbain. En effet, les villes offrent le meilleur comme le pire des environnements pour la santé et le bien-être. D'un quartier à l'autre, les influences positives et négatives se cumulent et distribuent inégalement les facteurs de risque qui déterminent le bien-être et la santé des populations. L'objectif de cette thèse de doctorat est d'analyser les inégalités spatiales de santé à Dakar. Une telle analyse requiert des données géoréférencées de qualité. Cependant, les sources de données existantes sont souvent imparfaites, peu désagrégées, et peu reliées aux facteurs contextuel tels que l'environnement de vie et la qualité de l'habitat. Cela a été démontré dans l'évaluation de la qualité des sources de données état civil et DHIS 2. Les recensements s'avèrent toutefois très utiles. La base de données du recensement de 2013 est la principale source de données utilisée pour mener cette recherche. Sa spatialisation s'est faite via un système d'information géographique (SIG). Des données d'occupation et d'utilisation du sol, extraites de l'imagerie satellitaire à très haute résolution ont également été utilisées, de même que des données d'enquête sur l'infection au COVID-19. Dans une approche de géographie de la santé, cette thèse de recherche développe une analyse systémique qui inscrit les états de santé dans un cycle, un processus d'interactions entre les populations, leurs habitats et les différents comportements qui créent et modifient les composantes de l'habitat. Les différences d'états de santé résultent de ces processus d'interactions. Il est à signaler que la question des échelles devient alors cruciale dans la mesure et la compréhension de ces interactions, en raison des risques de confusion entre corrélations statistiques et causalités réelles. Les indicateurs utilisés pour refléter les états de santé sont le taux de mortalité et le taux d'incidence du COVID-19. Les facteurs de risques contextuels pouvant jouer un rôle de déterminants sont la densité de population (mesurée à l'échelle du quartier et à l'échelle du bâti), l'assainissement, l'accès à une source d'eau saine, la pauvreté, l'éducation, la qualité de l'habitat, la physionomie de l'espace résidentiel (planifié, non-planifié) et l'état de l'environnement (bâti dense, moins dense, la végétation et la présence d'eau).

Une typologie des quartiers combinant données de recensement et variables extraites de l'imagerie satellitaires a d'abord montré le caractère hétérogène de Dakar du point de vue de la qualité de l'habitat. Quatre profils de quartiers ont été identifiés : les quartiers spontanés, les quartiers spontanés moins denses, les quartiers résidentiels avec espaces administratifs, commerciaux et de services, et les quartiers résidentiels. L'habitat spontané prédomine à Pikine, alors que le spontané moins dense est plus présent à l'est dans les aires périphériques de la ville. Très présent à l'ouest et au centre-ville, l'habitat résidentiel se diffuse aussi vers le front d'urbanisation à l'est. L'examen de la distribution du taux brut de mortalité selon les profils de quartiers identifiés montre une incidence du taux brut de mortalité plus élevé dans les quartiers spontanés, et plus faibles dans les quartiers résidentiels. Des tests statistiques confirment la significativité de ces différences. Mais le taux brut de mortalité s'avère trop général et sensible à la structure par âge de la population. Une décomposition du taux brut de mortalité en des taux bruts spécifiques s'est alors avérée nécessaire pour une lecture fine des

inégalités spatiales de décès. Les taux bruts spécifiques de mortalité qui ont été définis sont les suivants : enfants et adolescents (0-14 ans), adultes (15-59 ans) et personnes âgées (60 ans et plus). L'analyse de la distribution spatiale des taux bruts spécifiques de mortalité fait ressortir des clusters de surmortalité, aussi bien pour les enfants et adolescents, les adultes et les personnes âgées. Les clusters de forte mortalité adulte se retrouvent principalement dans les quartiers de Pikine, puis dans une moindre mesure à Rufisque et Guediawaye. Les clusters de faible mortalité adulte se retrouvent quant à eux dans les quartiers du département de Dakar. Les clusters de forte mortalité des enfants et adolescents sont aussi localisés dans les quartiers de Pikine et Rufisque, alors que l'intensité de la mortalité des enfants et adolescents est plus faible dans quelques quartiers du département de Dakar. Entre ces trois groupes d'âge, l'incidence de la mortalité est de loin la plus élevé chez les personnes âgées. Dans cette catégorie d'âge, les clusters de mortalité plus élevée se rencontrent d'abord dans les quartiers du centre-ville, puis en périphérie dans le département de Rufisque. Les clusters de faible mortalité pour les personnes âgées se retrouvent dans les quartiers de Pikine. Afin de tenir compte des effets de clusters, une régression géographiquement pondérée a été utilisée pour modéliser l'impact des facteurs de risque contextuels sur les variations spatiales des taux bruts spécifiques de mortalité. Les variables contextuelles associées aux variations spatiales de la mortalité sont d'abord la densité de population à l'échelle quartier, qui influence positivement la mortalité, quel que soit le groupe d'âge considéré. Ensuite, l'éducation, l'assainissement, l'accès à une eau source d'eau saine, la pauvreté, l'inaccessibilité aux services de santé par manque de ressources financières influencent également les taux de mortalité par âge, avec une intensité des relations qui varie toutefois d'un groupe à l'autre et selon les quartiers. Les facteurs de risque contextuels semblent jouer un rôle plus important chez les enfants, adolescents et adultes que chez les personnes âgées.

La distribution spatiale de l'infection au COVID-19 fait apparaître aussi l'existence de clusters. Ces clusters sont inégalement distribués dans l'espace avec une concentration de clusters de forte incidence du COVID-19 dans les quartiers du département de Dakar. La modélisation spatiale de type « machine learning », basée sur des arbres de régression boostées a permis de mesurer, sur l'échantillon d'individus infectés au COVID-19, le rôle de la densité de population à l'échelle du quartier et à l'échelle du bâti, ainsi que celui des espaces administratifs, commerciaux et de services, considérés comme les espaces les plus connectés. L'infection au COVID-19 subit en premier lieu l'influence de la densité de population à l'échelle du quartier, suivie de celle des espaces administratifs, commerciaux, et de services et enfin de la densité de population à l'échelle du bâti. La cartographie du risque d'infection au COVID-19 montre une probabilité d'infection plus élevée dans la partie Ouest, particulièrement dans les zones de forte densité de population et abritant les activités administratives, commerciales et de services. Cette modélisation spatiale de l'infection au COVID-19 à l'échelle de l'individu a été comparée à une modélisation à l'échelle agrégée (quartier). Les indicateurs de performance mettent en exergue une plus grande robustesse des analyses spatiales au niveau individuel. Mais ces résultats sont à manier avec prudence en raison des risques d'erreurs écologiques. Les erreurs écologiques surviennent lorsque des facteurs de risque collectés à l'échelle agrégée servent à inférer sur des phénomènes analysés à l'échelle individuelle. Si les corrélations statistiques sont aisées à établir, les causalités réelles le sont beaucoup moins pour des phénomènes qui sont d'abord sociaux, avec leur lot d'incertitudes. En effet, même si les quartiers diffèrent en fonction du statut économique de leurs habitants, les citadins pauvres ne vivent pas tous dans des taudis ou dans des zones concentrant des désavantages sanitaires, et ceux qui vivent dans des taudis ou dans des zones concentrant des désavantages sanitaires ne sont pas tous pauvres. La complexité est telle qu'il s'avère impossible de tirer des conclusions sur les causes directes de la mortalité, en se basant seulement sur l'influence des facteurs de risque contextuels. Malgré tout, ces tendances sur les variations spatiales d'état de santé à Dakar et les facteurs contextuels qui leur sont associées constituent des résultats de recherche originaux dans la planification urbaine et les interventions de santé publique.

Une des contributions scientifiques majeures de cette thèse est la décomposition du taux brut de mortalité en des taux bruts spécifiques. Une autre contribution scientifique importante est la modélisation spatiale tenant compte de l'autocorrélation spatiale et de la sur-dispersion des données. Cette thèse de doctorat a par ailleurs démontré la faisabilité des analyses spatiales à travers les données du recensement pour pallier l'imperfection des autres sources de données. Les données de recensement offrent surtout une plus large couverture spatiale, ce qui permet de s'extraire des analyses fragmentées et peu reliées aux facteurs de risque contextuels en milieu urbain. Les données issues de l'imagerie satellitaire à très haute résolution ont également beaucoup contribué à l'identification des facteurs de risques contextuels. Cependant, les problèmes d'autocorrélation spatiale et de sur-dispersion mentionnés ci-dessus, ont tout de même rendu les modèles de régression classiques inadéquats. Si la tendance selon laquelle les espaces les plus proches tendent à présenter des valeurs similaires (loi fondamentale de la géographie) fait de l'autocorrélation spatiale une caractéristique intrinsèque des données géographiques, la sur-dispersion est quant à elle liée aux données, caractérisées par exemple par une résolution temporelle limitée, qui fait des décès un événement rare et mène à une proportion élevée de petits nombres dans la base de données du recensement. De ce constat émane une recommandation forte, celle d'améliorer le système d'état civil en tant qu'outil de veille sanitaire, et de rendre les données existantes beaucoup plus accessibles et exploitables. Cela permettra de tirer profit de la meilleure manière des techniques de modélisation spatiale, qui pour pouvoir fournir des résultats robustes, requièrent des données en grande quantité et de grande qualité.

Table of contents

Chapter 1

Introduction

This introduction presents the state of the art of health inequalities in African urban areas and the determinants of geographical health inequalities. It then presents the methods and tools of spatial analysis, presents the research objectives, the research hypotheses and the structure of the thesis. The health situation in Dakar, which is our case study, will also be presented as well as an overview of the research methodology.

1.1 **Health inequalities in African urban areas**

Health inequalities in African urban areas are to a large extent the result of an urbanisation process that is both rapid and largely unplanned. While Africa is the least urbanised continent, it is also the most rapidly urbanising region in the world. The proportion of the population living in urban areas in Africa almost doubled over the last 50 years, from 22.6% in 1970 to 43.5% in 2020 (United Nations 2019). Estimates also suggest that half of the African population will be urban by 2035, and this proportion will have reached 58.9% in 2050 (United Nations 2019). However, this rapid urbanisation is not supported by an equivalent economic dynamism. Described as recent, brutal and massive, it is distinct from current trends of urbanisation in other parts of the world (Boyce et al. 2019). In Africa, high rates of urbanisation combined with limited economic progress lead to a growing concern with deteriorating health conditions (United Nations 2020).

In general, urban life is attractive because cities offer unique opportunities to residents to increase their income and benefit from education, health and social services. Although on average, urban dwellers benefit more from these services than rural ones, their distribution is uneven across urban areas (Dye 2008). African cities are highly heterogeneous, especially because the rapid and unplanned urbanisation increased the proportion of the population living in slums and informal settlements. Health inequalities are an unacknowledged feature of urban life, mirrored by social inequalities (Utzinger et Keiser 2006; Alirol et al. 2011). Many African cities face a double burden of sustained high rates of infectious and chronic diseases (Hussein 2014; de-Graft Aikins et al. 2010). As a result, mortality rates show unique aspects in their levels and trends.

1.1.1 Communicable diseases

In the era of epidemiological transition, where infectious diseases are declining and chronic diseases increasing, Africa still suffers from the magnitude and persistence of infectious diseases. Since the last decade, the share of deaths caused by infectious diseases in Africa continuously decreased, from 61.4% of total deaths in 2010 to 52.9% in 2019 (WHO-GHE 2019). Despite this steady decline, infectious diseases are still responsible for over half of deaths in Africa and remain a major threat to public health and economy. In fact, the absolute number of deaths caused by certain diseases still increases, an aspect which is hidden by proportions due to population growth. Deaths due to tuberculosis, for example, increased in absolute numbers from 347 310 in 2000 to 395 734 in 2015. At the same time, it relatively decreased from 52.6% to 40.2% (WHO-GHE 2019).

In five of the six WHO regions in the world, the burden of non-communicable diseases exceeds that of communicable diseases. The only exception is Africa where infectious diseases remain the major cause of death (WHO-GHE 2019). Table 1.1 reports the five dominant infectious diseases in Africa in term of deaths in 2019.

Table 1.1: Five major communicable diseases in Africa

Cause of deaths	**Number of deaths**	**Share of total deaths**
Lower respiratory tract infections	774 252	9.9
HIV/AIDS	434 543	5.6
Diarrhoeal diseases	496 278	6.4
Malaria	388 229	5.0
Tuberculosis	378 193	4.9

Source: WHO, 2019.

Taken together, the above five major infectious diseases are responsible for more than half (59.9%) of infectious diseases' deaths and for 31.8% of deaths from all causes, all age-groups combined. Over time, the downward trend in the share of deaths due to infectious diseases concerns only lower respiratory tract infections, HIV/AIDS, and diarrhoeal diseases. Since 2015, malaria has overtaken tuberculosis in the ranking of infectious diseases deaths in Africa. Indeed, deaths caused by malaria and tuberculosis increased between 2016 and 2019, respectively from 4.6% to 5.0%, and 4.6% to 4.9% (WHO-GHE 2019). As the leading cause of deaths in Africa, lower respiratory tract infectious diseases affect people's airways and lung. They often originate from viruses and bacteria, and occasionally from fungi and parasites. The most common lower respiratory tract infections are bronchitis and pneumonia. Tuberculosis is not included in lower respiratory tract illnesses because it can infect almost any part of the body, and is a major opportunistic disease that disproportionately affects people living with HIV. Among the diarrhoeal diseases, cholera, that is caused by bacterial infection, is the most common one, which is characterised by three or more loose or liquid stools a day, i.e. more than is considered normal for a person (www.africacdc.org). Out of the 15 countries across the world that account for 80% of the malaria burden, 14 of them are Sub-Saharan African countries (WHO - WMR 2017).

In addition to these endemic communicable diseases, the African region is also prone to infectious disease outbreaks due to cholera, measles, yellow fever and more recently COVID-19 (Mboussou et al. 2019; Lone et Ahmad 2020). Although they often emerge outside cities, infectious diseases have high potential to reach urban areas and to spread rapidly across cities due to the highly mobile urban population (Alirol et al. 2011), and higher population densities (Hazarie et al. 2021). COVID-19 is a noteworthy example showing the vulnerability of African urban areas to the risk of importation of infectious diseases, as well as their potential to emerge and spread quickly (Gilbert et al. 2020; Shrestha et al. 2020). While it emerged outside the African continent, the COVID-19 disease moved rapidly to Africa through urban centres, especially in dense and well-connected capital cities. The introduction of COVID-19 in Africa also illustrated the role of air transportation for disease diffusion, in the context of a globalised world (Gilbert et al. 2020).

1.1.2 Non-communicable diseases

Non-communicable diseases increased in low- and middle-income countries over the last two decades and are becoming an important contributor to the disease burden also in Africa. They accounted for 37.1% deaths in 2019, an increase of 3.6% compared to 2015 (WHO-GHE 2019). The share of deaths due to non-communicable diseases is growing rapidly on the continent, and by 2030, they are projected to exceed deaths due to communicable, maternal, perinatal and nutritional diseases combined. The five major non-communicable diseases in 2019 in Africa are displayed in Table 1.2.

Table 1.2: Five major non-communicable diseases in Africa

Cause of deaths	**Number of deaths**	**Share of total deaths**
Malignant neoplasms (Cancers)	534 293	6.9
Ischaemic heart diseases	429 179	5.5
Strokes	425 719	5.5
Cirrhosis of the liver	195 320	2.5
Diabetes	177 079	2.3

Source: WHO, 2019.

Malignant neoplasms refer to the family of various types of cancers and since 2010, they have overtaken stroke and remain the most important killer among chronic diseases. Stroke and ischaemic heart diseases are the main cardiovascular diseases that affect people in Africa.

Health related aspects of urbanisation are more perceptible when it comes to risk factors associated with non-communicable diseases. Tobacco use, physical inactivity, harmful use of alcohol, and unhealthy diets are particularly recognised as risk behaviours provoking chronic diseases, and are closely linked to a changing lifestyle driven by urbanisation (Juma et al. 2019). Characteristic for chronic diseases is their long duration, further related to an improved life expectancy. They are expected to be relatively higher for richer households in urban residential neighbourhoods.

The COVID-19 pandemic revealed that people living with chronic diseases such as hypertension and diabetes are more exposed to severe cases of COVID-19 infections (Sanyaolu et al. 2020). For example, 61% of the COVID-19 patients in hospitals in South Africa suffered from hypertension and 52% from diabetes. With almost half of all COVID-19 cases and deaths registered on the continent, South Africa is among the most affected by the pandemic in Africa (OCHA 2020 ; Stiegler et Bouchard 2020). These facts provide evidence that communicable and non-communicable diseases are linked. Indeed, the aetiology of diseases goes beyond the mere interference of infectious diseases, as shown above for tuberculosis and HIV/AIDS. A chronic illness can render a person susceptible to infectious diseases, whilst several infections play a role in the emergence of non-communicable diseases (Ogoina and Onyemelukwe 2009). This raises the issue of the complexity of the causal links in the mechanisms of disease transmission, with differential exposure of socioeconomic groups and different roles of the environment.

1.1.3 Mortality and life expectancy

Monitoring the level and trends of mortality is important since it directly reflects the health status of the population. Mortality rate declined since 2000 from 13 to 9 deaths per thousand in 2010 (Table 1.3). This value continued to decline over the last decade to 7 deaths per thousand in 2019 (WHO-GHE 2019). The annual percent of decrease was 0.7% between 2000 – 2010 and 0.3% between 2010 – 2019, respectively (WHO-GHE 2019). Despite this overall decline, Sub-Saharan Africa has the highest mortality rate for children under 5 (0-5 years), older children and adolescents (5-15 years), and young adults (15-24 years) compared to the rest of the world (Masquelier et al. 2021; 2018). Along with the declining overall mortality, life expectancy has risen at the same time. In Sub-Saharan Africa for example, it increased from 50.45 years in 2000 to 61.62 years in 2019 (UN-Population Division 2019). Life expectancy is another key metric to assess population health.

Table 1.3: Mortality trends from 2000-2019 in Africa

Year	**Deaths**	**Population**	**Rate (per 1000)**
2000	8 696 864	660 221 000	13
2010	8 059 944	858 413 000	9
2015	7 813 407	982 485 000	8
2019	7 786 394	1 091 759 000	7

Source: WHO 2019.

This general trend at the continental level may vary when looking more closely at the situation within countries and between sub-population groups. As illustrated by Günther and Harttgen (2012) in a longitudinal analysis in SSA, inequalities of mortalities within cities and subgroups of city dwellers is more pronounced than the overall country-level differences and the subnational rural-urban inequalities. For example, child mortality rate is higher in urban slums, compared to urban formal settlements. Still worse, child mortality rate in urban slum areas is higher than in rural settlements, for 3 out of 20 African cities. Analysing the evolution of mortality rates at the intra-urban level also provides an indirect gauge of the effect of the rapid and unplanned urbanisation in Africa.

Most importantly, the occurrence and spread of epidemic infectious diseases in African cities hinder the shift towards stage 3 of the theoretical epidemiological transition model as defined since 1971 by A. Omran. The three major successive stages of this model are (1) the age of pestilence and famine, (2) the age of receding pandemics, and (3) the age of degenerative and man-made diseases (Omran 2005). The HIV/AIDS pandemic has long been a pitfall in this transition (Caselli 2001). This is still ongoing, and the actual disrupt caused by COVID-19 is further slowing down the transition.

The crude mortality rate measures the frequency of deaths in a defined population during a specific time interval (Jacqueline 2009), as calculated by following formula:

$$\text{Crude mortality rate: } \frac{Deaths\ occuring\ during\ a\ given\ time\ period}{size\ of\ the\ population\ among\ which\ the\ deaths\ occurred} \text{ x } 10^{n}$$

The crude mortality rate is frequently used as a measure of mortality. Other common measures are standardised and specific mortality rates. While the crude mortality rate combines deaths from all causes for the whole population, specific mortality rates disaggregate the crude mortality rate by age, sex, or according to socio professional groups. The most frequent specific mortality rates are infant mortality rate, neonatal mortality rate, postneonatal mortality rate and maternal mortality rate. Further, the crude mortality rate has a limited explanatory power since it neglects differences in the age structure of a defined population. For example, a region with an elderly population may have a relatively high crude mortality simply because of its age-structure. To eliminate this age effect, the standardised mortality rate is used. The age standardised death/mortality rate obtained by the direct standardisation method is the rate that would be observed for a defined population that has the same age structure as the reference population. It is calculated by weighting the age-specific mortality rates observed in the defined population by the age structure of the reference population. Finally, different crude mortality rates can be used, with deaths and the corresponding population size broken down by age-group (Leclerc 1990; Bernard et Lapointe 1987).

1.2 Geographical determinants of health inequalities: the role of contextual factors

From the above, it seems obvious that the health status within African urban areas is heterogeneous. It is therefore of high interest to study health risks in African urban areas by taking into account the spatial heterogeneity of living conditions. Health indicators and their determinants must be disaggregated at finer levels to better grasp the urban reality. Indeed, the social and health conditions of the poorest city dwellers are not only comparable but sometimes even worse than those of rural populations.

Much of the literature has shown that medical aspects are not sufficient to predict the determinants of mortality and morbidity inequalities in dense and heterogeneous African urban areas (Soura 2009; Salem 1998a). Additionally to biological factors, such as age of mother, gender, birth spacing and birthweight, non-medical conditions also play a crucial role in explaining differences in health outcomes (Iyun 1993). City dwellers suffer disproportionately from poor health and these inequalities can be attributed to differences in social and environmental living conditions (WHO-UN HABITAT 2010).

This section introduces the role of contextual factors in population health, a topic that is not new to research. Already in 1854, John Snow tracked cholera deaths by water districts to explain why certain areas of London had higher mortality rates. He found that cholera contamination and deaths were related to unsafe water consumption in poor districts, and identified the responsible pumps (Whitehead, 1869). In African urban environments, Salem (1998) mentioned the unfavourable urban environment of neighbourhoods with excess mortality in Pikine, a suburb of Dakar.

Several concepts describe these characteristics (Mosley et Chen 1984; Emch et al. 2017) and the present thesis particularly refers to the one by Emch et al. (2017), known as the triangle of human ecology. It describes how human behaviour, in its cultural and socioeconomic contexts, interacts with environmental conditions to produce or prevent disease among susceptible people (Emch et al. 2017). This constitutes the aetiology, or the causal evolution of health and disease. Human disease and health vary across the earth’s surface, and these variations are the result of the interaction between population, habitat and behaviour, as outlined in Figure 1.1.

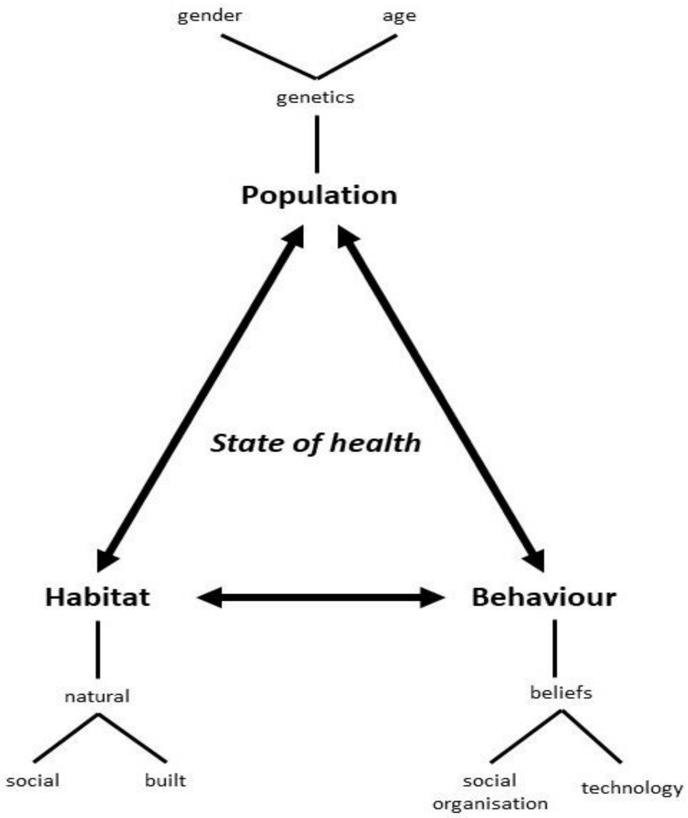


Figure 1.1: The triangle of human ecology (Emch et al. 2017, Health and medical geography).

The state of health of a given community is the result of interactions between these three broad categories, each of them also having its own process. More formally, people, through their behaviour, create habitat conditions, expose themselves to or protect themselves from habitat conditions, and move from place to place. The triangle of human ecology is very useful when investigating causal pathways of disease or trying to identify opportunities for interventions or prevention of diseases and poor health (Emch et al. 2017).

1.2.1 Population

Population is a core determinant of health in the sense that it refers primarily to the biological organism of individuals as a potential host for disease. Individual demographic and economic characteristics are important aspects of this vertex largely influenced by behaviour, under certain habitat conditions. The triangle of human health emphasises age, gender and genetics as fundamental population characteristics related to health. Age is a key factor for the health status since several types of diseases – chronic or infectious – put certain population groups at risk based on their age (Thomas, 2018). In addition, the ability of a population to cope with various kinds of threats depends, in some ways, on the stage of life. Accordingly, the mortality of children under 5 years has for long gained more attention in targeted interventions, because demographers and public health experts consider it as a critical issue and a core indicator of the development of families, societies, and the world at large (Yaya, Bishwajit, et al. 2018). More broadly, life tables are an essential tool to monitor age-specific demographic and health trends. The use of child mortality as a proxy for adult mortality seems ambiguous, especially due the variations and changes in health determinants.

The individual's ability to cope with threats also depends on the genetic susceptibility or resistance, its nutritional, immunological, and physiological status as well as its metabolism (Emch et al. 2017). Throughout life, these characteristics change with age. For example, children and women (especially during pregnancy) require special attention regarding infectious diseases such as malaria, while adults are suspected to have an immune system that recognizes and copes with a wider variety of infectious agents. Ageing leads to exposure to long-term chronic and degenerative diseases. Nevertheless, also younger individuals can be burdened by chronic diseases because familial genetic characteristics may favour asthma and diabetes for example (Smew et al. 2020).

1.2.2 Habitat

Habitat quality plays a crucial role when intending to make cities a healthier place. Three types of habitats are defined: the natural, the built and the social. The natural habitat addresses the environmental effect on human health, the weather and climate that we experience, the availability of water or food, the plant and animal species that surround us (Emch et al. 2017). Known as ecological factors, natural habitat is also a synonym for land cover types, excluding built-up areas. Their influence on health may be either direct or indirect. In Sub Saharan African urban environments, Boyce et al. (2019) analysed the direct effects of ecological factors on health by the emergence or re-emergence of infectious diseases, caused by the degradation of ecosystems, the intensification of agriculture and the increased opportunities for the human-animal interface, especially with rodent or peri-domestic species serving as reservoir for diseases. Other effects of ecological factors on mortality can be drought causing malnutrition due to famine, or rainfall favouring the development of water related-diseases (Dos Santos et Henry 2008; Henry et Dos Santos 2013). The latter poses a notable risk in the city environment, particularly in slums and shantytowns (Alirol et al. 2011).

The built environment (commonly referred as "built-up areas" in land cover classifications) refers to living houses and other buildings in which people spend a significant amount of time. It is often the first man-made landscape component that stimulates well-being and health. Some basic differences between slums and formal settlements illustrate this point as a house that is constructed with solid materials differs from a dwelling built from materials at hand. Thatched roofs offer nesting opportunities to insects, while roofing material of concrete slab and corrugated iron exclude them. Barriers for insect vectors, the location of a kitchen (inside or outside a dwelling because of smoke), and the presence of piped water, flush toilets, and an improved sanitation system are decisive (Emch et al. 2017). Several studies focused on patterns of the built environment (as a proxy) to relate the socioeconomic status to health conditions (Bawah et Zuberi, 2004). In the study of the ecology of urban malaria in Africa, increased built-up density reduced the risk of malaria transmission (Kabaria et al. 2016), which may be explained by the absence of water bodies and vegetation providing breeding sites for Anopheles mosquitoes within such densely built-up areas (Trape et al. 1992). Indeed, people living closer to natural surroundings may be more exposed to such vector-borne diseases than households living farther away.

Land use patterns can also significantly impact on health. If land cover – defined by the attributes of the land surface and immediate subsurface – refers to the natural and built environment, land use as a component of habitat concerns the purpose of human exploitation of the land (Lambin et Geist 2008). Geographers look within settlements at the spatial arrangement of residences, or more explicitly, land uses. These spatial arrangements cause specific patterns that influence living conditions and thus, the health status. Planned residential areas are characterized by linear settlements, where houses are lined up and spaced along straight roads. Conversely, spontaneous or deprived settlements result in houses that are sprawled over the land surface, without land planning that could promote various facilities, including sanitation systems, a solid waste disposal and health services. When located at the coast or river margins, spontaneous settlements are vulnerable to hazards such as landslides and floods. Humans create much of their disease environment (Emch et al. 2017).

The social environment is composed of the groups, relationships and societies in which people live, in the sense that there are social interactions between them that influence the health and behaviour. The concept of neighbourhood as a social entity emphasizes the importance of affinities, i.e. the existence of strong social relationships between residents of this social entity. When they struggle together to improve their physical environment, exchange information on health practices and risks, as well as on the availability of healthcare and the way to overcome health and safety shortcomings, they build social capital. The role of social capital is further examined in the analysis of health behaviour. If health conditions are shaped by population habitat, individual and collective behaviour determine the causal relationships, in other words the variations of opportunities or barriers to health.

1.2.3 Behaviour

Behaviour is essential in the triangle of human ecology, and refers to the processes by which individuals create, modify and adapt to habitat conditions (natural, built and social habitats). These processes also result in movement from one location to another, both for humans and other elements of pathological systems (e.g. pathogens, vectors and hosts). Beliefs, social organization and technology are important aspects of behaviour (Emch et al. 2017). They either expose to or protect individuals and populations from hazards. Density, a primary health risk factor, reflects the interaction process of the human heath triangle, as it is part of the population vertex, but related to behaviour leading to crowded or dispersed, scattered groups of individuals in a given habitat. To illustrate clearly the idea of health behaviour, we quote the statement of Emch et al. (2017), when analysing research results of Mackenbacket (2001). The latter had found that mortality was about 50% higher in the poorest neighbourhoods (with a higher percentage of unemployed and poor persons) compared to the wealthiest ones (with a lower percentage of unemployed and poor persons). Emch et al. (2017) stated that: "After all, just because you're poor doesn't mean you are biologically less fit than a person who is wealthy. Thus, there must be something about living in poor communities that affects people's health and behaviours, ability and willingness to utilize health care, or stress levels." The contextual effects largely (but not exclusively) include how socioeconomic status and social capital act as determinants of health. Several researchers measured the influence of socioeconomic factors (education, income and consumption, housing characteristics, households assets) on health conditions (Pongou et al. 2006; Fotso et Kuate-Defo 2005; Harttgen et al. 2020; Tusting et al. 2017; Doctor 2004; Bawah et Zuberi, 2004; Balk et al. 2004). Households' income and consumption are major components that reflect socioeconomic effects on health (Sharif et al. 1993). Higher income typically leads to better health by enabling the household to spend more on healthcare. It also leads to improved housing, better sanitation and therefore reduces the exposure to disease. With higher income and consumption, nutritional intake and balance is likely to improve, resulting in greater resistance to disease. Education is closely related to the socioeconomic position and is also included as element of behaviour affecting health. Higher education of mothers and the head of household in general contributes to beliefs and attitudes that promote health. In Ghana, Kofie et al. (2008) associated buruli and ulcer and diarrheal diseases to inadequate access to safe water, unsanitary conditions, inappropriate waste disposal methods, lack of education and information as well as restricted access to health care services. Social capital refers to the benefits derived from the connections between neighbours. As a product of the social environment, social capital influences health behaviour through social contagion, collective efficacy and informal social control (Emch et al. 2017). In this respect, higher literacy rate of women at community level is a meaningful contextual effect. Indeed, educated mothers can influence the health behaviour of their non-educated counterparts, diminishing the risk of child death. This is a social contagion. Health services are an integral part of the human habitat, but shaped of elements of behaviour, such as technologies and social organisation. For example, the availability of health care services can create a contextual effect for persons living in the same area (Soura, 2009). Collective efficacy and informal social control are the strategies used by communities to fill the gap of access to health care services. When they combined efforts to upgrade existing or build new facilities, the improvements will translate into lower morbidity and mortality.

1.3 Spatial epidemiology

Spatial epidemiology is a growing discipline, at the crossroads of health geography and epidemiology. It analyses the spatial variations in disease risk and incidence. For the ultimate goal to assist public health, the following questions should be addressed (Chen et al. 2014):

(a) What are the distributions of diseases across space and how do they interact with their/the? environment? What are disease origins, destinations, and spreading channels?
(b) What are the potential spreading patterns of a disease regarding the potential habitat of its host and its environment?
(c) Which diseases will successfully spread around the globe via global travelling and trading as well as wildlife movement?
(d) Which parts of regions hold the greatest risk of disease exposure given urban and regional host habitats and population distributions as well as intercity and regional transportation networks?
(e) Which population groups are the most vulnerable to a disease?

Below, the concepts of epidemiology and health geography are briefly introduced, followed by the description of methods and tools used in spatial epidemiology.

1.3.1 Spatial epidemiology and health geography

Spatial epidemiology emerged in a context where epidemiological studies were limited to the identification of groups at risk. The main objective of epidemiology is to describe and measure the characteristics of a disease or health condition in a population, to estimate the risk, and to help search for the causes of this disease or health condition. Epidemiology identifies relationships between diseases and risk factors in order to reveal their mechanisms and processes. Disease risk factors in epidemiology include for example the cholesterol level, proportion of smokers (and non-smokers), and obesity for cardiovascular diseases, pulmonary infection, and diabetes. From a public health perspective, the added value of spatial epidemiology is the spatial analysis and prediction of the outcome of a disease over time and across space. Although spatial epidemiology and health geography fill the gaps of epidemiological studies by adding at-risk areas to at-risk groups, the area of interest of health geography is larger and includes fields that are not covered by spatial epidemiology (e.g. social aspects of health, health behaviour, healthcare systems). Health geography emphasizes the sense of location, a concept that shifts health issues to specificities such as the geography of mortality, or the analysis of the spatial and social distribution of mortality (Salem, 1998a), the geography of diseases (analysis of the spatial and social distribution of diseases), the geography of the health system (location of establishments, analysis of the spatial distribution of the health system, spatial disparities in the health system, accessibilities, inequalities, studies of flows, use of health services, models of hospital attractiveness), the geography of populations and territories in relation to health (health assessment, vulnerabilities, planning, allocation of resources, influences of health in the geographical construction of territories), the regional health planning (identification of needs and priority objectives, forecasts, definition of health basins and health territories). It also addresses public health problems related to behaviour and infrastructures (traffic accidents) (Souris, 2019). The territory, a space of multifactorial construction, is the core of this approach, defining health geography as the study of the quality of the population health, its behaviour and factors of its environment that contribute to the promotion or deterioration of its health (Picheral, 1898). In other words, it is concerned to discover the combinations of a set of factors that, within a given area, differentially expose populations to certain health risks, and that distribute the populations exposed to such a risk unevenly in space (Salem, 1993). Ultimately, place is a living construct of importance, and the purpose of health geography is to examine spatial processes that underlie states of health in different places (differentials in vulnerabilities and exposures), and variations in spatial patterns and causal links of the determinants.

1.3.2 Spatial statistical analysis and modelling

Spatial statistical models involve the statistical analysis and modelling of a health outcome variable with its locations and its potential impacting factors in space and time domains (Chen, 2014). Health outcome variables, either mortality rate or disease incidence rate, often do not follow a Gaussian distribution and are not independent of the development of the statistical methods. other than mathematical models, spatial statistical models have the advantage of depicting regional risk factors and incorporate both, spatial and temporal residual variation in the analysis (Chen, 2014). Thus, they include advanced analysis techniques that enable to move beyond the conventional statistical methods by incorporating spatial effects (spatial dependency or clustering) or dealing with spatial data problems (Lawson, 2013).

In spatial statistical analysis and modelling, three broad classes of methods are usually applied: disease or mortality mapping, cluster analysis and ecological analysis (Chen, 2014). The first class concerns the use of models to describe the overall distribution of a disease or the mortality on the map, with the purpose of mapping area-specific relative risk estimates. Standardized incidence ratio (SIR) and standardised mortality ratio (SMR) are the two commonly used relative risk estimates for disease and mortality data, respectively (Moraga, 2019). There are three main categories of spatial data: (1) geostatistical data that are primarily parametrized with continuous values and chosen locations; (2) aggregated lattice data based either on a regular or irregular lattice; and (3) point process data containing observations with responses for the random spatial process (Cressie, 2015). Areal or aggregated lattice data result when a fixed domain is partitioned into a finite number of subareas in which outcomes are aggregated (Moraga, 2019). They form the basic health information data. Number of malaria cases per neighbourhood, number of household deaths over the last 6 months or the proportion of people living in poverty in census tracts are examples of areal data. Various mapping techniques have been developed and are applied for each category. They range from spatial smooth mapping techniques to spatial interpolation methods (Chen, 2014). Spatial smooth mapping techniques are used to clean the noise, reduce the abrupt risk variation in point process disease data, and estimate disease risk by computing the value at a location as the average of its nearby locations (Lawson, 2013). Spatial interpolation methods, including kernel and krigeage estimations, construct risk trend surfaces over the entire study area, based on data collected from limited locations (Berke, 2004).

The second class, the cluster analysis, is used to reveal or detect spatial autocorrelation, that is the unusual concentration or non-randomness of the health outcome variable in the space and time domain. Clustering can be tested globally or locally and remains a useful resource in public health measures. The most common measures of global and local spatial autocorrelation are Moran's and Geary's index, as well as kuldorff spatial scan statistics.

The last class, ecological analysis is used to analyse the relationship between the spatial distribution of a health outcome variable and measured risks factors that can potentially impact the occurrence of the health outcome variable at an aggregated spatial level. Spatial regression is used for ecological analysis, to model spatial effects. There are several types of spatial regression models that use a spatial weighting matrix to account for spatial autocorrelation existing in variables and residuals. Starting with classical regression model in equation 1:

$$y = X\beta + \varepsilon \tag{1}$$

where y is the health outcome variable for neighbourhood, $X\beta$ is the matrix of a set of explanatory variables with their coefficients β, and ε is a normally distributed error term. In a classical ordinary least squares (OLS) regression, ε should display a normal Gaussian distribution with the mean 0. When spatial autocorrelation exists in the health outcome variable, the residual plot of the error will no longer be normally distributed. It highlights a clustering effect, and the relationship between the health outcome and explanatory variables is no longer constant over the study area. Known as non-stationarity of the regression coefficients, spatial autocorrelation leads to biased estimates, with an under estimation of the variance of the value of estimated coefficients.

To account for a clustering effect, a spatial error model uses a spatial contiguity matrix to incorporate the spatial configuration in a regression model as in equation 2:

$$\begin{gathered} y = X\beta + \mu \\ \mu = \rho W + \varepsilon \end{gathered} \tag{2}$$

where W is the contiguity matrix, and β and ρ are parameters to be estimated in the model. The parameter matrix ρ indicates the extent to which the variation of Y can be explained by the neighbouring values.

Another spatial regression model, called spatial autoregressive model (SAR), uses the spatial weight matrix as shown in Equation 3:

$$y = \rho W y + X\beta + \varepsilon \tag{3}$$

This model is similar to the lagged dependent variables for time series regressions.

More advanced models can be derived from the combination of equations 2 and 3 to build other geographically weighted regression models. The spatial error model in equation 2 accounts for spatial autocorrelation in the way that the residual ε might affect (or be related to) the residuals in the neighbouring regions. In equation 3, the spatial autoregressive model conceptualizes spatial relationships by the value of y in a given area that might impact on (or be related to) the value of y in a neighbouring region. This depicts the spill over effects across the whole area. The spatial autoregressive model can also be conceptualized by the value of X in a given area that might affect (or be related to) the value of y in a neighbouring region. This model is called lagged independent variable and is formulated as follows:

$$y = X\beta + WX\theta + \varepsilon \qquad (4)$$

These models are flexible regarding the format of the input data and the selection of parameters. However, they are challenging because of the demands on the data. They often require a great effort to preprocess the data in large quantity and variety. Missing data, underreporting, uncertainty, and zero-count data often occur in mortality and disease data. Overdispersion is a particular issue in the spatial modelling of health outcome variables due to the scarcity of health data on a small and temporal geographical scale.

The advent of machine learning methods gives an opportunity to deal with data problems in spatial modelling. In particular, spatial modelling techniques based on boosted regression tree models is useful to deal with spatial autocorrelation and fit complex interactions. Cluster analysis can be used as a preliminary analysis to give insights about the use or non-use of spatial regression models in ecological analysis. Although its strengths in revealing the variation of regression coefficients across the study area, spatial regression of aggregated spatial units raise the issue of ecological fallacy when interpreting the results.

1.3.3 GIS and remote sensing

Spatial statistical analyses largely rely on geolocated data. The development of concepts, methods and techniques related to geographic information system has allowed the use of spatial analysis in spatial epidemiology (Souris, 2019). In the era of big data, GIS evolved and matured quickly to combine multi-sources and various format of spatial data (Zhou et al. 2020). At the same time, remote sensing, through very-high resolution satellite imageries, provides more detailed environmental and socioeconomic information for a better understanding of health-related determinants (Jia et Stein 2017; Georganos et al. 2019a). The COVID-19 pandemic has shown the noteworthy benefit of GIS techniques in disease surveillance studies. Zhou et al (2020) described the contribution of such techniques in the attempt to slow down the rapid spread of the coronavirus in China. The GIS techniques implemented allowed rapid visualization of epidemic information, spatial tracking of confirmed cases, prediction of regional transmission, spatial segmentation of the epidemic risk and prevention level, as well as solid spatial information support for decision-making, formulation of measures, and assessment of effectiveness of COVID-19 prevention and control. Similar approaches were developed for African countries in order to evaluate their preparedness and vulnerability regarding the risk of importation of COVID-19 (Gilbert et al. 2020). The use of remote sensing significantly increased since the 1990s and provides many opportunities for spatial epidemiological studies. Spatial epidemiologists often refer to remote sensing in infectious diseases research as a tool to assess a greater range of environmental factors that promote disease transmission, vector production, and the emergence and maintenance of disease foci, as well as the risk for human-vector contact (Beck et al. 2000). In African urban environments, environmental variables derived from remote sensing have been extensively used to map the transmission of malaria and associated risks factors (Dlamini et al. 2019; Georganos et al. 2020; Kabaria et al. 2016; Machault et al. 2010). Medium resolution images, such as Landsat and SPOT, formerly provided information on vegetation cover, landscape structure and water bodies. But recent earth-observing missions with much higher spatial and spectral resolution allow to measure additional environmental information. Very-high resolution imageries can now provide socioeconomic information through spatial metrics. The landscape metrics include for example planned or deprived areas and can distinguish vegetation into natural and agricultural areas (Tais Grippa et Stefanos Georganos 2018). Thus, such very-high resolution remote sensing data sources offer new opportunities in ecological studies. However, they contain inherent uncertainties and require technical skills and thoughtful preprocessing steps for accurate assessment and interpretation.

In summary, these methods and tools are suitable and highly valuable for a health geography approach. Spatial modelling provides original techniques to analyse contextual effects of mortality and morbidity by taking spatial effects into account.

1.3.4 The issue of scale of analysis: why does it matters

Individuals are part of a neighbourhood, a community and a region, a nested spatial hierarchy where the elements described above and their mechanisms take place (Figure 1.2).

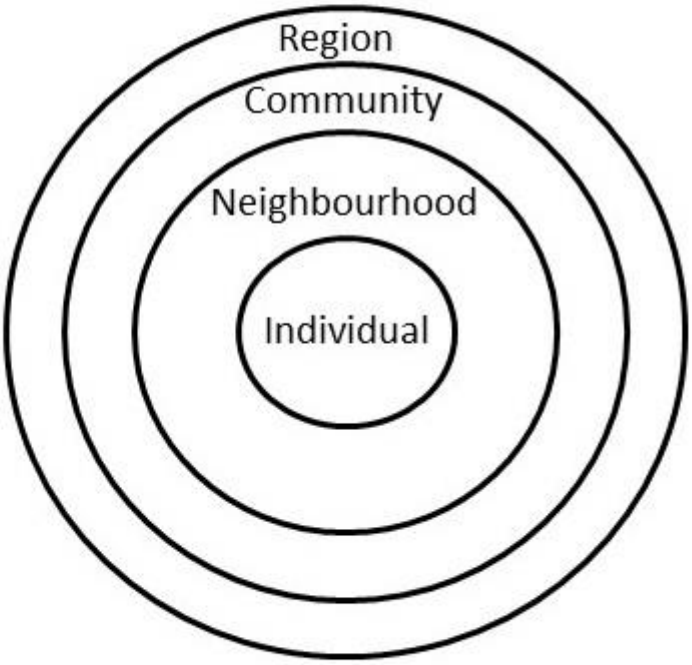


Figure 1.2: The nested spatial hierarchy, (Emch et al. 2017, Health and medical geography).

The question of the scale of analysis is crucial when measuring the effects of contextual risk factors (Salem 1998a; Soura, 2009). The same characteristics offer different views depending on the scale of measurement. In light of such complexity, there is consensus that the finer the spatial scale, the better the understanding of reality. In the United Nations report on health inequalities, the authors insist on the need to disaggregate the information at finer levels to better reveal urban reality (WHO-UN HABITAT 2010). Many researchers agree on this principle. Salem (1998) however states, that a health phenomenon is not always best described at the finest spatial scale, but that justified spatial aggregations may be meaningful by providing different information, and that the objects of research change according to the scale of analysis. In any case, there is a need to include group- or macro level variables in epidemiological studies, and the selection of the appropriate contextual units and contextual variables is challenging (Diez-Roux, 1998).

To illustrate the complexity of the concept, one of the key elements in epidemiology, population density (expressed as the ratio of population per unit area) is a good example. There is no doubt that population density influences disease transmission. However, it should be measured at a spatial scale that is relevant to the health problem and to the population concerned: density of houses for diseases related to domestic constriction (communicable, such as measles for children, tuberculosis for adults and children, or non-communicable, such as burns from cooking or heating stoves);

density of the housing block for diseases related to peri-domestic constriction (communicable, such as helminthiasis, or non-communicable, such as falls); neighbourhood density for pathologies linked to the neighbourhood's living environment (transmissible, such as malaria, where density is a factor that rather diminishes exposure through dilution of anopheles bites, or non-transmissible, such as traffic accidents, or even arterial hypertension, which is undoubtedly favoured by the stresses of high densities). Although these three scales of density are linked, their interdependence is not necessarily identical throughout the landscape: one can find high domestic densities but low neighbourhood densities if there are not many buildings, and vice versa. There are no general rules for the scale and the level of detail of the object studied: depending on the accessible data and the objective of a study, very detailed or coarse analyses can be carried out at different scales (Salem, 1998a). Bhadra's (2021) mentions that population density calculated per total surface of an administrative area assumes a certain uniformity of the landscape which leads to some bias. For example, the inconsistent findings of different studies on the role of population density in the transmission of COVID-19 demonstrate that the concept of the scale of analysis is critical to examine the geographic determinants of health. In various contexts and at different scales of measurements, correlations between COVID-19 infections and population density have been found to be either positive or negative (Hamidi et al. 2020; Bhadra et al. 2021; Kadi et Khelfaoui 2020; Carozzi 2020). It remains a persisting challenge that correlation analyses and statistical associations are scale dependent. Known as the modifiable areal unit problem (MAUP), results can vary depending on the level of aggregation, which may lead to ecological fallacy. Although some methodologies exist to reduce the risk of ecological fallacy (Tuson et al. 2020), the choice of the scale of analysis is critical in the study of relationships between states of health and related determinants. Things become even more complicated when human behaviour is involved. While correlations are easy to establish, causalities may be doubtful due to mediation, moderation, simultaneity and unobserved heterogeneity effects (Table 1.4).

Table 1.4: Theoretical overview of the variety of links between a variable X and a variable Y

Mechanism	**Illustration**	**Interpretation**
Direct causal link	X ⟶ Y	X influences Y
Mediation effect	X ⟶ Z ⟶ Y	The effect of X on Y is mediated by Z
Moderating effect	X ↓ X ⟶ Y	The effect of X on Y is moderated by Z
Unobserved factor	X ↙ ↘ X - - - - - - Y	The variable Z influences both X and Y: there is no casualty between X and Y

Bidirectional causality	X ⇄ Y	X is both explanatory of- and explained by- Y

Source: Manuel de géographie quantitative, 2019.

1.4 Objectives and structure of the thesis

1.4.1 Objectives

Cities offer the best and the worst environment for health and well-being. The health status of urban city dwellers is affected by multiple determinants, and whether the urban influence is positive or negative largely depends on the place where one lives in the city. Within this framework, the general objective of this thesis is to analyse spatial inequalities of health in Dakar. This general objective is broken down into four specific objectives:

- Assess the potential of vital statistics and DHIS 2 data for spatial analysis
- Identify neighbourhood housing quality profiles and their relationship with crude mortality rates
- Understand the relationship between age-specific mortality rates (child-adolescent, adult and old age) and identified contextual risk factors
- Analyse the spatial distribution of COVID-19 infections.

In a health geography perspective, we develop a systemic approach, based on the triangle of human health, to study the geography of mortality and morbidity in Dakar.

1.4.2 Research hypothesis

The general research hypothesis underlying this work is the following:

Inequalities in states of health in Dakar are the consequence of contextual factors, including spatial variations in density, land cover and land use, as well as spatial inequalities in housing quality.

This general hypothesis includes four specific hypotheses:

- Available data are of sufficient quality to analyse health inequalities in Dakar
- Dakar is characterised by health inequalities that are to large extent related to inequalities in living conditions
- Determinants of mortality vary both across space and across age-groups

1.4.3 Thesis structure

After this introduction, the present thesis is divided into four parts: The first presents the existing health data sources in Dakar and discusses the motivation for choosing one or another for a special analysis of the geography of mortality; the second focuses on inequalities in housing quality in order to unravel the spatial heterogeneities of living conditions in Dakar; the third looks at mortality by age-group, analysing its geographical distribution and related determinants. The fourth is devoted to the spatial modelling of COVID-19 infection.

Chapter 2 provides the basis for the easy use of reliable mortality data in the spatial analyses conducted in this thesis. Health data were obtained from the national census bureau, civil registration office, DHIS platform, and surveys including DHS, verbal autopsy, GEOCOVID, and malaria surveys. In this thesis, census data are the fundamental data used to derive mortality indicators. Civil registration data are known to be a more complete source. This chapter evaluates the completeness of the data stemming from the Dakar's main civil registration office, and their potential for the spatial analyses carried out here. The potential of the DHIS 2 data is also examined.

Chapter 3 gives a detailed overview of the spatial heterogeneity of the living conditions in Dakar. A typology of neighbourhood was carried out by combining census and remote sensing variables, which allowed to derive neighbourhood housing quality profiles. These neighbourhood profiles range from spontaneous settlements, spontaneous low dense settlements, residential areas with administrative, commercial and services areas, and residential areas. The ultimate goal of this typology was to examine the impact of living conditions on health. Distributions and hypotheses were statistically analysed to describe the distribution of crude mortality rate across neighbourhood profiles.

Chapter 4 and 5 focus on spatial modelling. First, determinants of the geographic distribution of age specific death rates are modelled. A stepwise quasi Poisson regression model is performed, followed by a geographically weighted regression, which takes into account the presence of spatial autocorrelation in the spatial distribution of age specific death rates at neighbourhood level. Second, the spread of the COVID-19 infection in Dakar is modelled with a cluster analysis, the role of contextual risk factors is measured and areas at-risk of COVID-19 infection are mapped. An individually-based boosted regression tree model, and an aggregated-based negative binomial regression model were used and their performance compared. Figure 1.3 outlines the thesis structure.

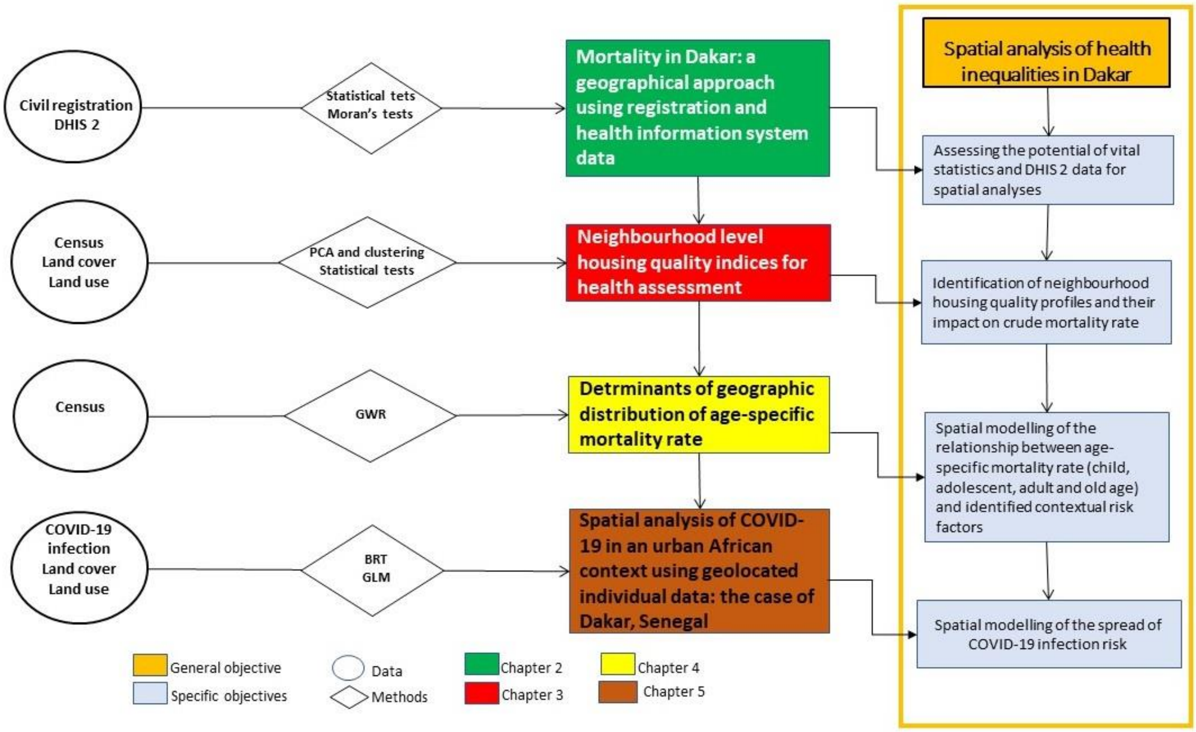


Figure 1.3: Overview of the thesis structure.

1.5 Case study: Dakar

Dakar is the capital city of Senegal, situated on the Cape Verde peninsula (Figure 1.4). The Dakar region covers an area of 550 km² located between longitude 17° 10 and 17° 32 West and latitude 14° 53 and 14° 35 North. In 2013, the population of Dakar was 3 137 196, which corresponds to almost a quarter of the whole population of Senegal (23.2%) on only 0.28% of the national territory. The age structure of the population shows that the Dakar region is very young; the under-20s represent 44.5%.

The Dakar region leads all other regions of the country in terms of population, economy and facilities. Since June 1958, the region has become the government headquarter by becoming the country's capital. As a result, it concentrates most of the industrial sector, commercial establishments and financial institutions. Dakar is the least poor region in Senegal, with a poverty index of 12% at household level. About 50% of monetary income is concentrated in Dakar. Households derive most of their income from the informal sector (27%), the private sector (24%) and transfers (12%). Dakar is also the region with the best health infrastructure and personnel in the country. Eleven of its twelve health facilities hold level 3 (i.e. health structures at the top of the health pyramid), and only here, all medical specialties are offered.

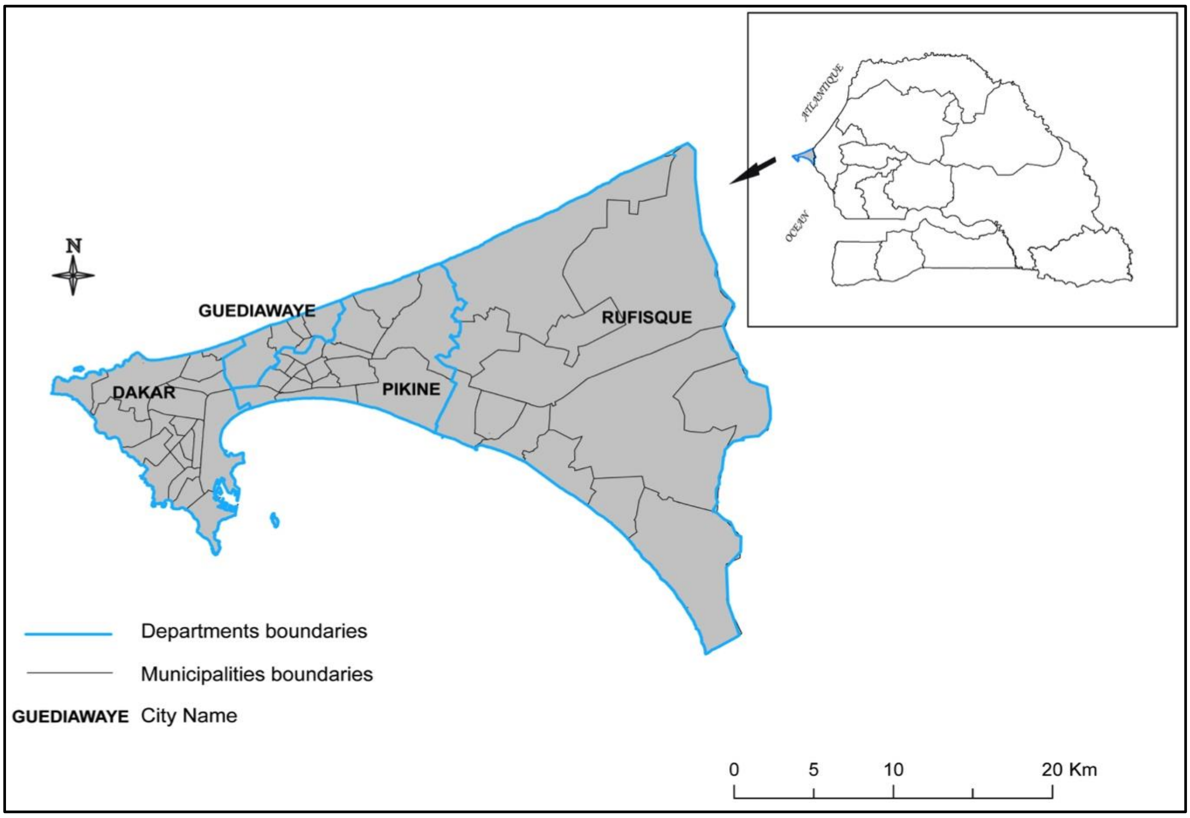


Figure 1.4: Location map of Dakar.

1.5.1 Health situation

Overall, Dakar is still doing better than other regions concerning the health situation. According to the last census, life expectancy at birth reached 69.9 years, compared to the national average of 65 years. The crude mortality rate shows an annual average of 5 deaths per 1 000 inhabitants, while the national annual average is 8 deaths per 1 000 individuals. Compared to other regions, Dakar combines the highest life expectancy and the lowest crude death rate. However, these global figures conceal very large intra-urban differences. The region is marked by significant inequalities with regard to access to health care services (Duboz et al. 2015). These are associated with socioeconomic inequalities: while eastern neighbourhoods are characterized by a low standard of living and have limited access to health services, western neighbourhoods have a higher living standards combined with better access to health care services (Ndonky et al. 2015).

1.5.2 The triangle of human health applied in Dakar

Based on the triangle of human health, a systemic approach for the analysis of mortality and morbidity inequalities in Dakar is developed in this thesis, which helps to illustrate the role of contextual factors on health. The triangle of human health in Dakar can be adapted to the framework of this thesis as follows (Figure 1.5):

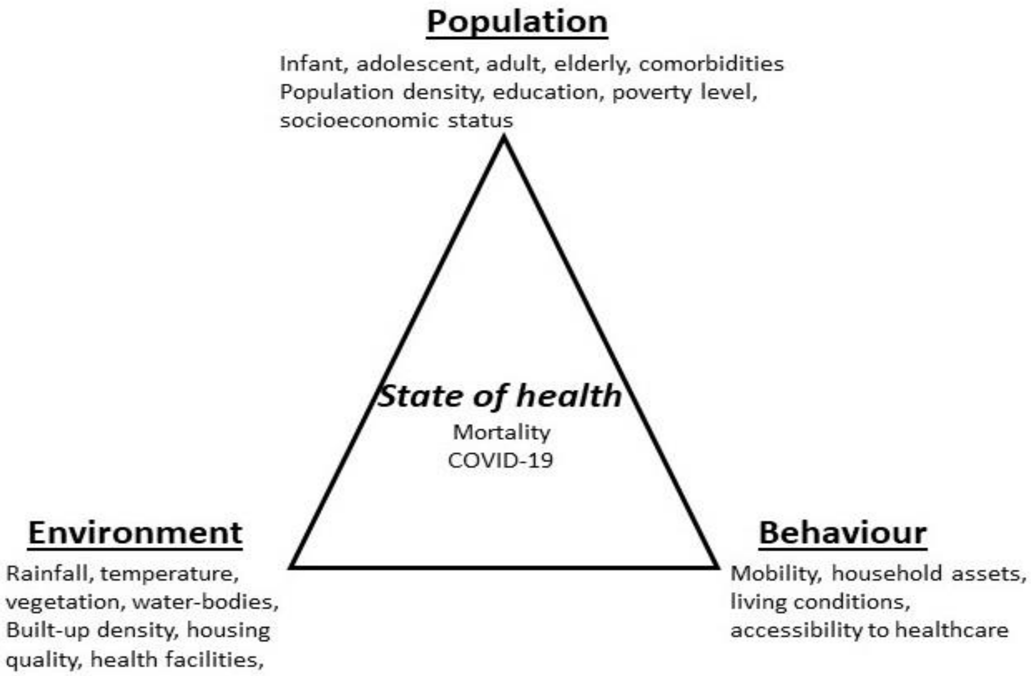


Figure 1.5: : The triangle of human ecology adapted to the framework of this thesis (following Emch et al. 2017)

Existing research on mortality and morbidity in Dakar remains scarce. The pioneer work of Salem (1998) highlighted the extreme urban heterogeneity and revealed that access to healthcare is the decisive variable correlating with mortality. He further identified high housing density and poor access to safe water sources as the main risk factors for mortality from parasitic diseases. However, due to the absence of age-specific mortality data, this study is limited to the analysis of crude mortality which is known to be confounded by population age structure. In addition, it only covers the suburb of Pikine, one of the four departments that compose the region of Dakar.

More recent studies focussed on particular diseases in Dakar: vector-borne diseases (Dos Santos et al. 2015), including malaria (Diallo et al. 2012; Machault et al. 2010), diabetes (Duboz et al. 2015), ischaemic and haemorragic events (Sagui et al. 2005), and influenza (K. Ndiaye et al. 2000). However, these studies are often very specific, non-spatialized case studies, and poorly related to contextual characteristics.

A detailed analysis of the wealth and poverty inequalities highlighted a marked distinction between a rich city centre and a poor suburb in Dakar (Borderon et al. 2014), which serve as a proxy for inequalities in access to health care (Ndonky et al. 2015). This contrast between the city centre and the suburbs is however gradually diminishing (Borderon et al. 2014; Ndonky et al. 2015). This analysis of wealth and poverty inequalities has already increased the understanding of inequalities in living conditions. Nevertheless, differences in living conditions still need to be linked to differences in states of health, and this at a detailed spatial level, using spatial analysis and modelling methods.

This thesis ought to fill this gap. It uses a health geography approach, linking states of health to a set of contextual risk factors through spatial modelling methods that account for potential spatial heterogeneities of the health indicators analysed, and thus improves the understanding of the spatial variations of states of health. The underlying principle of this thesis follows Tobler's first law of geography: "Everything is related to everything else, but near thing are more related than distant things" (Tobler, 1970).

This research focuses on two aspects of health geography: geography of mortality and geography of diseases. Mortality, as a synthetic indicator commonly used to reflect state of health, is analysed using the 2013 census database including the crude mortality rate as well as age-specific mortality rates. Age-specific mortality rates used here include: child-adolescent (0-14 years), adult (15-59 years) and old age (60 years and over) mortality rates.

After the Dakar region became the epicentre of the coronavirus in Senegal, regional health authorities wanted to take strong measures to constrain the epidemic. Therefore, a survey on COVID-19 patients was conducted to identify their demographic and socioeconomic characteristics, activities, housing characteristics, and their movements prior to infection. This survey allowed to identify the profiles of COVID-19 patients to better define control strategies. this survey data is used to conduct the geography of morbidity via a spatial analysis of the COVID-19 infection.

This research is conducted at neighbourhood level, the finest administrative unit in Dakar. Variables related to contextual factors come from both, census and very-high resolution remote sensing data sources and include land cover and land use information, household information relative to housing characteristics and household assets, collective equipment and private equipment, living conditions, poverty levels and information on access to health care.

Chapter 2

Mortality in the city of Dakar: a geographical approach using civil registration and health information system data[1]

[1] Adapted from: Diène A.N., Gadiaga A.N., Linard C. (2021). Mortality in the city of Dakar: a geographical approach using civil registration and health information system data. *In preparation.*

The best source for mortality analysis is civil registration. Dakar has good coverage in terms of deaths reported to the civil registry, but the data are under-exploited. This chapter examines the potential of vital registration data for spatial analyses of mortality. The examination of the potential for spatial analysis also includes DHIS data. This review will lead to the inclusion or not of vital statistics and/or DHIS 2 data in the spatial analyses of mortality inequalities conducted in this thesis.

Abstract: This paper reviews the available sources of mortality data for the city of Dakar, Senegal. It explores the quality of these sources and their potential use for spatio-temporal analyses. Two data sources are particularly analysed: civil registration and health information systems. Based on a comparison of deaths recorded in 2016 and 2017 in the Main Civil Registry Centre (MCRC) and in District Health Information Software 2 (DHIS2), this study sheds light on the shortcomings in recording health information and discusses the relevance of those data as sources of epidemiological information in order to improve public health policies. Comparison of the aggregated data shows significant omissions in these two sources, which makes it difficult to estimate reliable crude mortality quotients. Civil registry data may be more amenable than DHIS 2 data for spatial analysis of mortality. Nevertheless, these data should be interpreted with great caution. There are important limitations, particularly in the recording of places and causes of death, which lead to quality, completeness and reliability issues when using them for spatial analyses of prevalence rates and causes of death. These findings highlight the limitations of these data sources to date and the challenges faced by the two information systems in terms of collection, coordination and dissemination.

Keywords: *mortality; intra-urban inequalities; geographical approach; data sources; Senegal.*

2.1. Introduction

Sub-Saharan African cities are facing rapid urbanization and social change with a strong impact on morbidity and mortality rates (Baragatti and al. 2009 ; Thiam and al. 2017). Due to the lack of regular population-based surveys, these health dynamics remain poorly documented and most of the time, they are tackled only with biomedical paradigms that which take little account of urban health features (Satterthwaite and al. 2019). Collecting and analysing spatial data for better epidemiological surveillance is one of the main challenges for health systems, especially in rapidly growing urban areas like the city of Dakar. In this context, a geographical approach can contribute to a better understanding of mortality inequalities and the risks of death. Spatial analysis of causes of death contributes to the identification of intra-urban health inequalities and the major public health problems (Soumah and al. 2019). However, monitoring dynamics and spatial inequalities in mortality, requires health information systems, which can provide reliable and disaggregated estimates (Masquelier B. and al. 2016 ; Tabutin and Masquelier 2017).

According to the World Health Organization (WHO), although significant progress has been made in morbidity and mortality data collection, in many countries the health information systems still do not yet provide statistics to monitor Sustainable Development Goals (SDGs) indicators related to health (WHO 2017). Registration of deaths and causes of death remain particularly challenging in the vast majority of countries in sub-Saharan Africa (UN/ECA 2017). To sort out these shortcomings, African countries have set up an Africa Program for Accelerated Improvement of Civil Registration and Vital Statistics (APAI-CRVS 2017).

In Senegal, several projects have been implemented to improve the production of Civil registry statistics necessary for planning and achieving development goals[2].The civil registration system in Senegal is considered to be among the best in Africa, but the data recorded are still l less used for a public health perspectives (Salem 1998 ; RS/ANSD 2015). The main reason is the completeness issues in health data records that reduce their usefulness. In the rapid assessment of the civil registration and cause-of-death system in Senegal, the completeness rate was estimated at 28 per cent (RS/MATCL/ANSD 2012). The mission of civil registration offices is actually limited to

[2] These initiatives were reinforced in 2004 with the establishment of the National Civil Status Centre (CNEC) by Decree No. 2004-427 of 14 April 2004, which established a central body to modernize, coordinate and monitor the activities of the 679 civil registry centres throughout Senegal (RS/MICL, 2004).

the production of administrative documents (birth, marriage and death certificates). Data have not been published officially since 1975 (Cantrelle 1997).

In 2014, the Senegalese Ministry of Health and Social Action (MSAS) implemented the District Health Information Software 2 (DHIS2) to strengthen the health information system in Senegal. The Health and Social Information System Division (DSISS), which is attached to the Research Planning and Statistics Directorate (DPRS), is responsible for coordinating the collection and dissemination of data. The aim of the implementation of this tool is to improve data collection, processing and dissemination from the healthcare system, in particular the one relating to morbidity and mortality (RS/MSAS 2016). According to the Ministry of Health, this platform should centralise and integrate all health and social data in order to provide up-to-date information that can support decision-making. These reforms and new tools thus aim to improve Senegal's health information system, which is made up from very heterogeneous sources. These sources are often inadequate for spatialised analyses at the desired geographic and social scales.

While two systems recording deaths co-exist (health information and civil registration), there is no connection between them. In the absence of adequate mechanisms for the transmission, sharing and processing of data, statistics are used on a segmented basis and on a scale that is not well suited to the production of decision-making information.

The lack of spatial analysis limited greatly the use of operational indicators, in particular the identification of areas with specific mortality profiles, at risk of infant, maternal or premature excess mortality. Few recent studies have been dedicated to socio-spatial inequalities in mortality, particularly in Senegalese urban settings. Such studies of geographical inequalities in mortality provide useful information about populations, areas and periods with high risks. However, these studies research need to be scaled-up so that they can help to improve risk assessment and the effectiveness of evidence-based health actions.

The main objective of this chapter is to analyse the quality of the data recorded by the two sources of epidemiological information, in order to inform public decision-making. The study aims also at assessing the completeness, spatial representativeness, relevance and usefulness of these two sources of data to the public health. This form of research is useful for health surveillance as well as for better development of projects to improve civil registration and the health information system.

2.2. Study area

The city of Dakar, capital of the Republic of Senegal, is one of the four Departments of the urban agglomeration of the Dakar Region, along with Pikine, Guédiawaye and Rufisque (Figure 2.1). It was occupied by 1,252,786 inhabitants in 2016, representing 35% of the regional population according to projections by the National Agency for Statistics and Demography (RS/ANSD 2016). The Region of Dakar is marked by sustained urban growth generated both by demography and by its political, administrative and economic functions (Diop 2007). Indeed, the city is densely populated with density of more than 5,000 inhabitants per km² (RS/ANSD 2019). As stated by several authors, Dakar is a heterogeneous, even heteroclite and polycentric space (Dramé 2006; Borderon et al. 2014). This heterogeneity has to be taken into account, as each sub-area may present a specific epidemiological profile (Sy 2006; Baragatti et al. 2009). The degree of urbanisation, population densities, level of equipment, access to healthcare systems are factors of inequality for both within and among neighbourhoods. Over the last decade, geographical researches have been carried out on socio-spatial inequalities and their impact on health in the Dakar region. However, these researches focused on transmission of risk factors for specific diseases such as diarrhoeal diseases (Sy I 2006), malaria (Ndonky 2011) or access to health care (Dramé 2006). The study on the geography of neighbourhood-level mortality by Salem (1993) is over two decades old., His study revealed that urbanization has important effects on health, with poor sanitation, density and unequal access of health care as important risk factor. The study highlighted the interest of a geographical, i.e. spatialized approach to civil registry data and the need to integrate them into the health information system. Similar research may help to identify mortality risks and morbidity/mortality relationships to strengthen the health care system. They become fundamental in a context of epidemiological and socio-sanitary transitions which remain poorly documented in sub-Saharan African cities (Niang-Diene et Salem 2015).

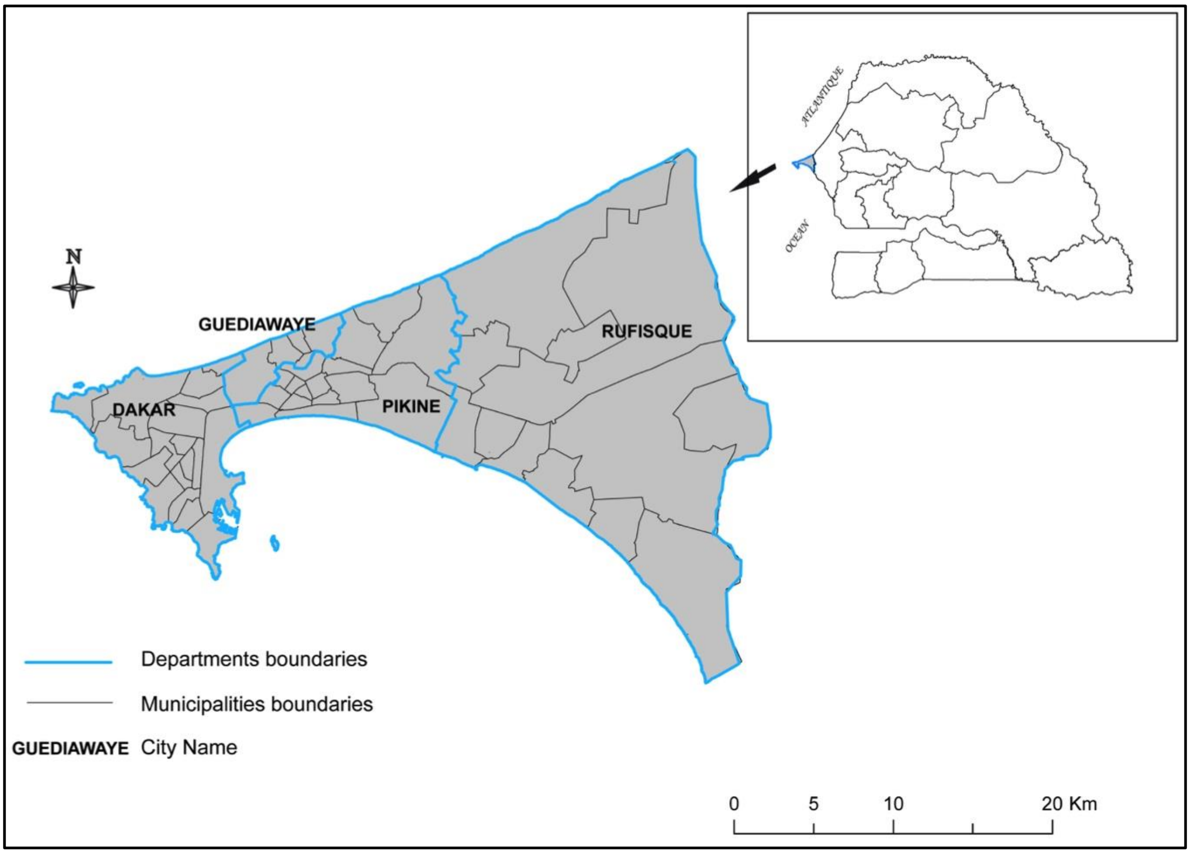


Figure 2.1: Administrative division of region of Dakar

The City/Department of Dakar is subdivided in 19 Municipalities called Communes[3] (Figure 2.1). These territorial authorities are headed by Mayors and they have a number of powers, including the management of neighbourhood shopping facilities, participation in household waste collection, maintenance of public streets and squares, and the administration of secondary civil registry centres (RS, 2013). These Secondary centres are institutionally attached to the main civil registration centre of the city of Dakar (MCRC).

[3] The commune is a local authority, a legal entity created by Decree. It brings together the inhabitants of the perimeter of the same locality composed of districts and/or villages united by good neighbourly relations and solidarity (RS, code of local authorities, 2013).

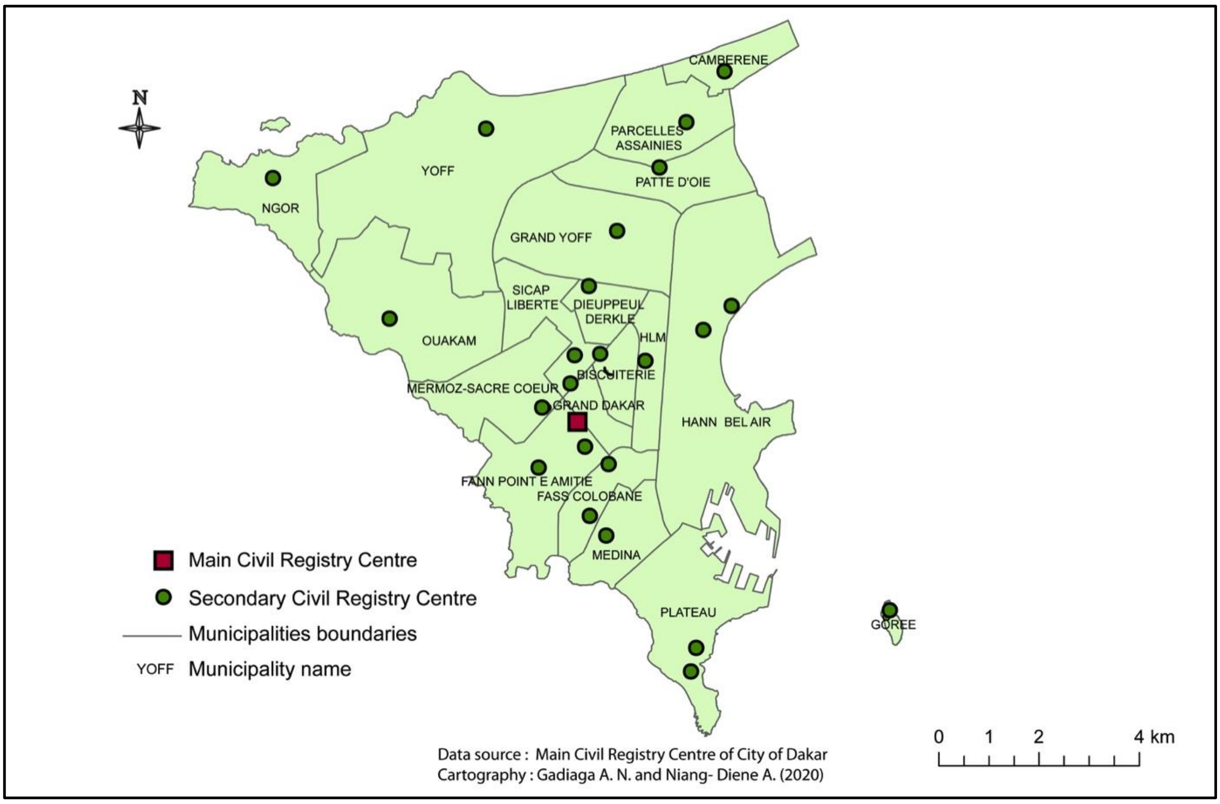


Figure 2.2: Location of the Civil Registry centres in the 19 Municipalities in the City of Dakar

The department of Dakar is divided into four health districts which are the health operational units in Senegal (Table 2.1 and Figure 2.2).

Table 2.1: Population and surface areas of the health districts of the city of Dakar (2017)

Health Districts	Number of municipalities	Total polpulation	Area (km^2)
Dakar-centre	7	366 259	13
Dakar-nord	4	493 337	22
Dakar-ouest	4	237 726	34
Dakar-sud	4	191 973	10
Total	**19**	**1 289 295**	**79**
Source : District Health Information Software 2 (2016-2017)			

It concentrates at least 40% of the country's health professionals and six of the eleven national hospitals (RS/MSAS 2019). This fairly dense health network nevertheless shows major inequalities in access to care for both public and private services (Dramé 2006b; Ndonky et al. 2015).

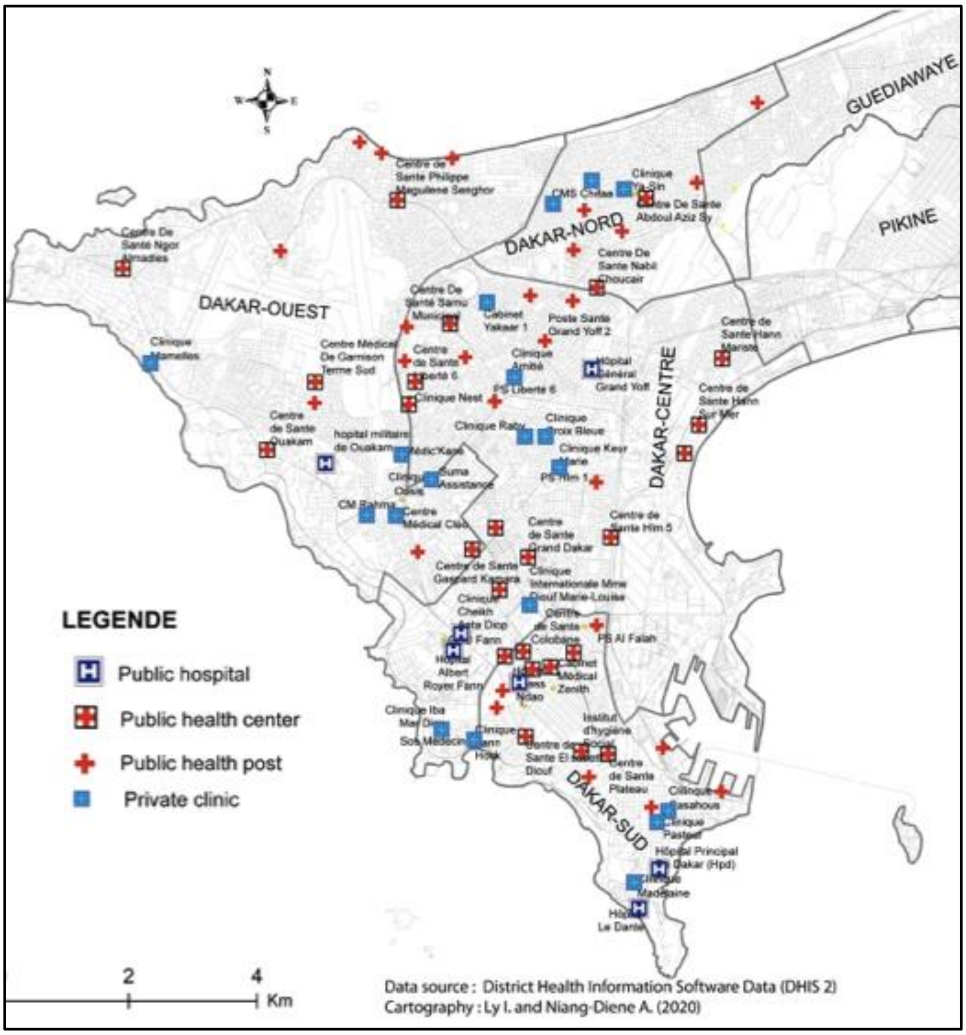


Figure 2.3: Public Health care facilities of the city of Dakar

2.3. Materials and methods

2.3.1. Data sources

The sources of data used in this article are the Civil Registry Centres in the city of Dakar and the District Health Information Software 2 (DHIS 2) of the Ministry of Health and Social Action statistics of 2016 and 2017. The data provided by the MCRC include data from 13 of the 24 civil registry offices in the city of Dakar (Table 2.2). While Secondary civil registry centres are supposed to transmit their data to the main office, half of them did not transmit any data to the MCRC for the years 2016 and 2017. In addition, data provided by Secondary Registry centres are limited to the total number of deaths per year. This aggregated data can only be used to estimate apparent volumes and crude mortality rates. Only deaths registered at the MCRC contain personal information of the deceased person (such as the place of residence) and can be used to analyse the spatial distribution of mortality. This spatial analysis can therefore only be carried out in the catchment area of the MCRC. This reality may be reflected in the representation of certain neighbourhoods of the city of Dakar in the statistics collected.

Table 2.2: Number of recorded deaths in the Civil Registry Centres (2016-2017)

Civil Registry Centres	2016	2017	Total
Centre Principal	1215	1395	2610
Fann-Point E-Amitié	150	147	297
Bourguiba - Biscuiterie	200	206	406
Dieuppleul-Derklé	231	300	531
Medina	450	500	950
Camberene	128	162	290
Yoff	311	398	709
Ngor	47	38	85
Ouakam	368	319	687
HLM	183	200	383
Hopital Abass Ndao	892	848	1740
Hann village	41	50	91
Hann sur mer	171	177	348
Mermoz Sacré-Cœur	-	-	
Patte d'oie	-	-	-
Parcelles assainies	-	-	-
Hopital Principal	-	-	-
Hopital Fann	-	-	-
Hopital Le Dantec	-	-	-
Plateau	-	-	
Fass-Colobane	-	-	-
Grand-Yoff	-	-	-
Sicap-Liberté	-	-	-
Gorée	-	-	-
Total	**4387**	**4740**	**9127**
Sources: Main Civil Registry Centre of the city of Dakar (2016-2017)			

The information collected at the MCRC includes:

- the number of registrations;
- age, sex and place of birth;
- place of residence;
- date, cause and place of death.

The District Health Information Software 2 (DHIS 2) is an online database that is supplied by the framework teams of the Medical Regions and Health Districts previously trained in the use of this platform (RS/MSAS 2016). It contains information on infrastructure, equipment, health personnel, as well as activities, morbidity and mortality recorded in public and private health care structures. Public health care facilities include hospitals, health centres and health posts. Private health care structures include clinics and doctors' offices. Access is granted to researchers who can benefit from an authorization to use the data from the dedicated site. The data used for this article concerns only deaths recorded in the four health districts of the city of Dakar (Table 2.3). It should be noted that this is only the population that was able to access to health services.

Table 2.3. Number of recorded deaths in the Health Districts of the City of Dakar

Health Districts	2016	2017	Total
Dakar-centre	415	382	797
Dakar-nord	346	1106	1452
Dakar-ouest	293	1521	1814
Dakar-sud	470	591	1061
Total	**1524**	**3600**	**5124**
Source : District Health Information Software 2 (2016-2017)			

In this database, the variables we were able to exploit include:

- age groups;
- sex;
- date (by month) of death;
- cause of death;
- and place of death.

Spatial information is limited to territorial health units and health facilities. The residential addresses of the deceased are not mentioned in the DHIS Data. Data from the 2013 national census (RPGHAE) and the 2016 and 2017 Demographic and Health Surveys (Continuous DHS) were used as references to compare and discuss vital statistics and DHIS 2. These surveys on the survival of the siblings of the women interviewed provided an estimate of maternal mortality and mortality of adults aged 15 to 50 years (RS/ANSD 2017). The levels, trends and differentials in perinatal, neonatal, infant and child mortality were also assessed in these DHS surveys.

2.3.2. Data auditing

In data auditing, we detected some errors related to data entry, such as erroneous or duplicated information. They have been found in the transcription of addresses, causes and places of death. To solve this issue, some modalities were grouped into categories and coded. For example, in the Civil Registry database, modalities such as "stillbirths fresh, stillbirths macerated, stillbirth n, stillbirth N, stillbirth N, ROM, HRP, MNF, late abortion" were grouped into a new variable called "stillbirths". Causes of death with numbers less than 3 have been added to the group of pathologies with which it is classified in the Ministry of Health nosological reports and in DHIS 2. Prior to the codification of addresses, we also corrected the transcription of the toponomy, for the variable place of residence of deceased, and thus harmonized neighbourhoods and municipality data inputs. It should be noted that this grouping of modalities, which facilitates data analysis, also leads to a loss of information on specific health problems (traffic accidents, meningitis, intoxications, etc.).

2.3.3. Cartography and spatial analysis

One of the objectives of this study is to locate and visualize deaths recorded at civil registration office (MCRC), at neighbourhood level. Spatial data on neighbourhood were obtained from the ASSESS project[4] (Improving death statistics to better monitor health trends in the Dakar) and the MCRC data were spatially joined with neighbourhood boundaries. As neighbourhood is the smallest administrative unit in urban areas, this cartography provides disaggregated data on deaths and their causes in Dakar. The spatial boundaries of the neighbourhoods were drawn on georeferenced maps in each municipality. This process resulted in the demarcation of 385 neighbourhoods in the 19 Municipalities (Fig. 4).

[4] www.assess-sn.org

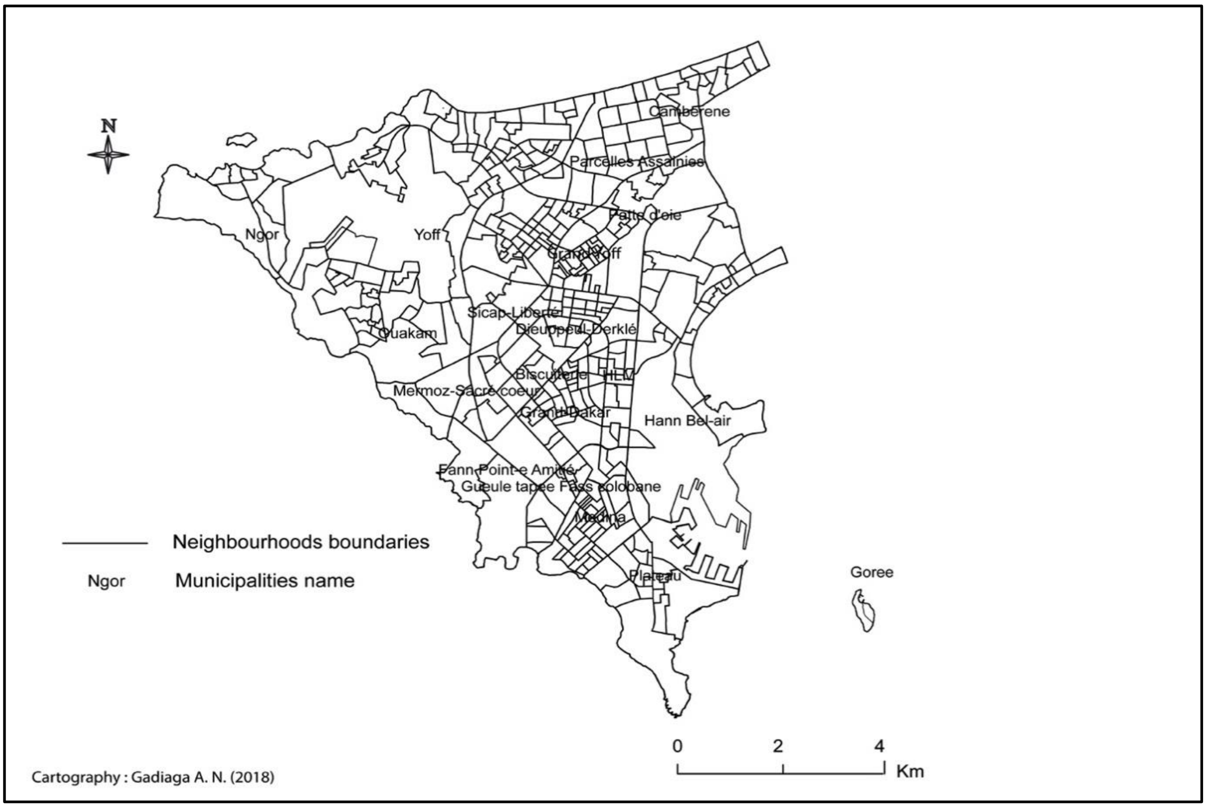


Figure 2.4: Neighbourhoods of the city of Dakar (2017)

MCRC statistics were subsequently codified and then associated with this cartographic database via a unique identifier (Syscol). With this data, we performed a spatial autocorrelation test, namely the Moran's index. The spatial autocorrelation of a quantitative variable refers to the similarity of its values according to their geographical location. It is positive when neighbouring geographical entities take close values for the variable under study. In the opposite case, the autocorrelation is said to be negative and then refers to a dissimilarity. The Moran's index I takes the following form:

$$I = \frac{n \sum_j \sum_j w_{ij}(Yi - \bar{Y})\,(Yj - \bar{Y})}{(\sum_{w_{ij}})\ \sum_i (Yi - \bar{Y})^2}$$

which Y denotes the variable under study and refers to the number of deaths by place of residence (neighbourhood). Similarity (or dissimilarity) also implies presence of clusters (neighbourhoods with excess mortality). The autocorrelation test was computed using R statistical software. At the same vein, we used package season of R statistical software to analyze seasonality of deaths.

2.4. Results

Aggregate data reveal important differences, particularly in terms of deaths recorded by the two sources (Table 2.4). The mortality recorded in the civil registration centres (9127 deaths) is higher than the DHIS 2 statistics (5124 deaths). This discrepancy can be explained by several factors, including better registration at the civil registry level, the missing data in DHIS 2 (particularly in the health districts of Dakar-ouest and Dakar-sud) and the proportion of deaths occurring at home. Further analysis is needed to measure the extent of these gaps and their spatial distribution.

Table 2.4: Number of recorded deaths (2016-2017)

Year	Main Civil Registry Centre (MCRC)	Secondary Civil Registry Centres	District Health Information (DHIS 2)
2016	1215	3172	1524
2017	1395	3345	3600
Total	**2610**	**6517**	**5124**
Sources: Main Civil Registry Centre of the city of Dakar (2016-2017) District Health Information Software 2 (2016-2017)			

It is possible to compare civil registration data with census data for the department of Dakar. The volume of deaths recorded in the civil registration offices represents a proportion of 75% and 78% of the total deaths reported in the census, respectively for the years 2016 and 2017.

2.4.1. Data completeness

Analysis of the volumes of deaths shows different completeness rates depending on variables (Table 2.5). As a whole, for each variable, the information is available for at least eight out of ten deaths, except for the causes of deaths (76 % and 70%, respectively in MCRC and DHIS 2), as well as the place of death in the civil registration data (28%). However, this completeness rate only indicates the quantity of data collected, not its quality. Data quality was assessed by taking into account the completion of the various items, which made it possible to identify the percentage of "missing" or "incomplete" data in a column. Important differences are also noted between the two databases. The date of notification of deaths is more accurate in the MCRC, where the day, month and year are mentioned. In the DHIS 2 data, the statistics are aggregated by month and by health care facility. In addition, the personal data of the deceased such as date and place of birth and place of residence are not taken into account on this platform. The lack of information on these variables is already a major limitation for the comparison of the data provided by these two sources, especially about geographical inequalities of mortality.

Table 2.5: Data completeness rate (in percentage)

Variables	Main Civil Registry Centre (MCRC)	District Health Information (DHIS 2)
Date of birth	82	not indicated
Place of birth	92	not indicated
Date of death	100	100
Cause of death	76	70
Place of death	28	100
Place of residence	99	not indicated
Age or age group	97	98
Gender	89	86
Municipality or District	96	100
Sources: Main Civil Registry Centre of the city of Dakar (2016-2017) RS/MSAS/District Health Information Software 2 (2016-2017)		

Another important limitation is the lack of information for the variable "deaths due to diseases and other causes" in the health districts of Dakar-ouest in 2016 and Dakar-sud in 2016 and 2017 (Table 2.6). These omissions reduce the completeness of the data provided by DHIS 2, which must be interpreted with caution.

Table 2.6: Number of recorded deaths in District Health information Software Data (DHIS 2)

Health Districts	2016				2017			
	Stillbirth	Neonatal mortality (0-11mois)	Maternal mortality	Diseases and other causes	Stillbirth	Neonatal mortality (0-11 mois)	Maternal mortality	Diseases and other causes
Dakar-centre	221	45	1	148	230	56	5	96
Dakar-nord	278	35	12	21	328	42	6	736
Dakar-ouest	228	59	6	-	208	81	9	1223
Dakar-sud	335	127	8	-	375	200	16	-
Total	**1062**	**266**	**27**	**169**	**1141**	**368**	**36**	**2055**
Source : District Health Information Software 2 (2016-2017)								

Any analysis of these official notifications must therefore take into account these limitations in the collection of data and their quantitative and qualitative shortcomings.

2.4.2. Distribution by age and gender

The classification of age groups in the DHIS2 by the Ministry of Health was also used for MCRC. Analysis of the distribution of deaths based on this typology shows that stillbirths and infant mortality are the most represented in the statistics collected (Table 2.7). Nevertheless, significant differences are noted depending on the two sources. Stillbirth corresponds to 14% of deaths registered by the MCRC, against 43% in the DHIS 2. Neonatal mortality (0-11 months) were the subject of 35% of notifications in the MCRC, compared to only 15% by the DHIS 2.

Table 2.7: Distribution of proportional mortality by age in 2016-2017

Age groups	Main Civil Registry Centre (MCRC)		District Health Information (DHIS 2)	
	Numbers	Percent (%)	Numbers	Percent (%)
Stillbirths	362	**14**	2203	**43**
0-11 months	919	**35**	807	**15**
12-59 months	261	**10**	372	**7**
5-14 years	418	**16**	383	**7**
15-19 years	104	**4**	166	**3**
20-25 years	130	**5**	208	**4**
26-49 years	131	**5**	486	**9**
50-59 years	104	**4**	147	**3**
Over 60 years	102	**4**	260	**5**
Undetermined	79	**3**	92	**2**
Total	**2610**	**100**	**5124**	**100**
Sources: Main Civil Registry Centre of the city of Dakar (2016-2017) District Health Information Software 2 (2016-2017)				

These differences can be explained firstly by the non-registration of deaths due to diseases and other causes in the health districts of Dakar-ouest in 2016 and Dakar-sud in 2016 and 2017. To mitigate this effect, data for these two districts were deducted and the proportions were re-estimated for the districts of Dakar-centre and Dakar-nord. The results show similar trends that seem to confirm a greater reporting of stillbirths and infant deaths, which represent between 10% and 15% and 30% and 80% respectively of the mortality reported in 2016 and 2017 by these two health districts (Fig. 5).

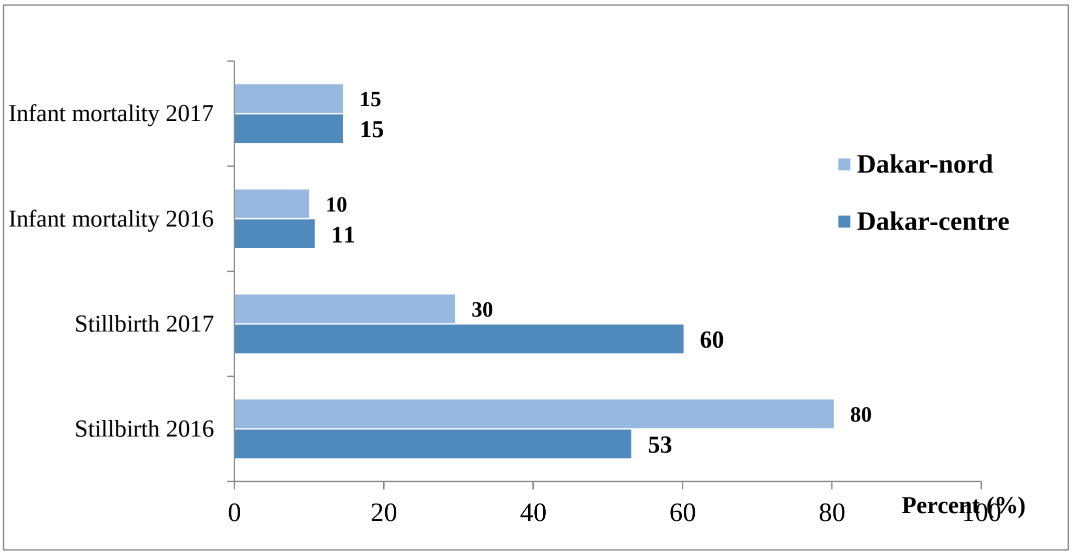


Figure 2.5: Distribution of proportional mortality in Dakar-centre

Source: District Health Information Software 2 (2016-2017)

Whatever the differences in proportions, stillbirths dominate the statistics in both sources. These results can be interpreted by various non-exclusive factors, including a priority given by the Ministry of Health to counting perinatal and infant mortality, which encourages health workers to provide better information on this variable. They may also be due to high perinatal mortality. In comparison, the results of the 2017 continuous DHS give a child mortality rate of 41.6 ‰ at the national level (RS/ANSD 2017) while those of the Multiple Indicator Cluster Survey (2015-2016) show an infant mortality rate of 32 ‰ for the Department of Dakar (RS/ANSD 2016b).

The distribution by age and gender structure seems to indicate a difference in stillbirth, which is reversed in both databases (Fig. 6 and Fig. 7). This raises questions about the registration modalities in both systems. However, in Senegal, the masculinity ratio at birth is 105 boys for 100 girls. A few singularities should be noted, including a greater number of death declarations for the adult male population at the civil registration. Surveys in the general population give similar results with mortality rates of 1.9 ‰ and 1.5 ‰ deaths, respectively for men and women aged 15-49 (RS/ANSD 2017). In the health system, the female population of reproductive age (15-59 years) is the most represented. This difference can be explained either by maternal mortality or by the greater use of health care services by the female population in these age groups (Dramé 2006a; Niang-Diène 2019).

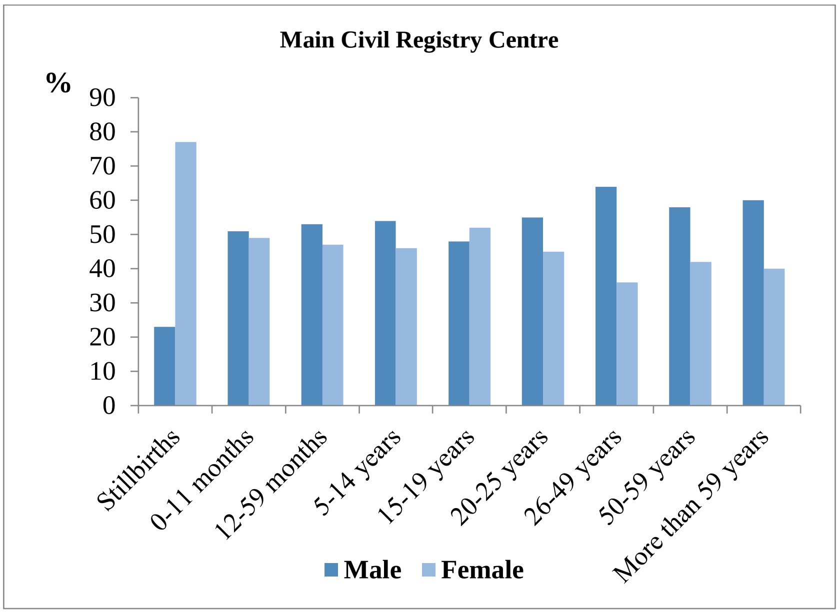


Figure 2.6: Distribution of deaths recorded by Main Civil Registry Centre by age and gender (in 2016-2017)

Source: Main Civil Registry Centre of the city of Dakar (2016-2017)

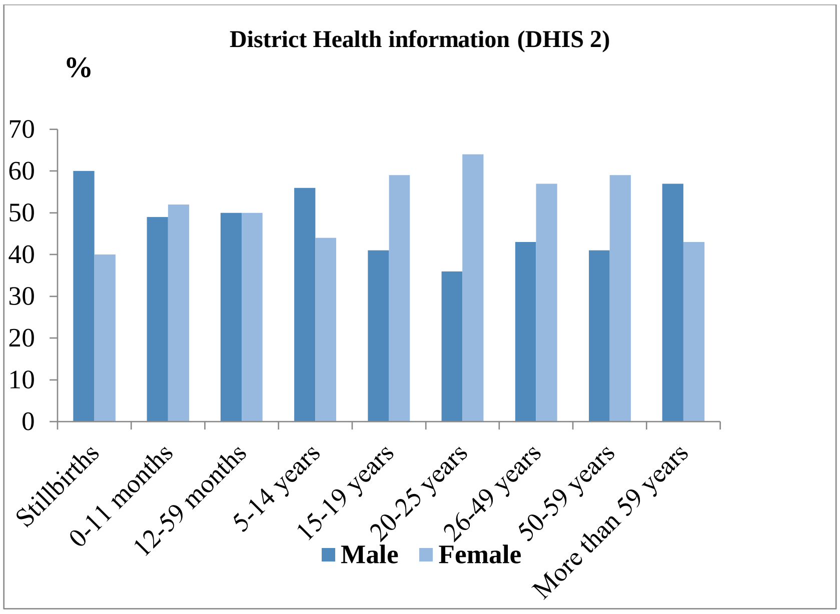


Figure 2.7: Distribution of deaths recorded In DHIS 2 by age and gender

(in 2016-2017)

Source : District Health Information Software 2 (2016-2017)

The maternal mortality ratio is estimated at 236 maternal deaths per 100,000 live births in the five years preceding the 2017 continuous DHS (RS/ANSD 2018). Differences in the reported mortality of adult males in the marital register compared to the health care system can also be explained by procedures for inheritance or for the establishment of administrative documents for descendants (Salem 1998).

2.4.3.Apparent rates and spatial distribution of mortality

The significant deficiencies already reported in the two databases do not allow for the estimation of reliable crude death rates by administrative unit either from the MCRC or from DHIS 2. Indeed, half of the secondary civil registration centres have not provided their data. Similarly, the Dakar-Ouest and Dakar-Sud health districts did not provide statistics on adult mortality in 2016.

Given these shortcomings, we mapped and analysed spatial distribution of deaths according to:

- residence of deceased ;
- place of death.

2.4.3.1. Reported place of residence: distribution by neighbourhood

As the District Health information statistics do not include addresses, spatial analysis was attempted from the MCRC database. In the latter, the overwhelming majority (80 %) of the addresses could be attached to municipalities and neighbourhoods of the region of Dakar. However, it is important to note the sizeable proportion of deaths (875 or 34 %) coming from other cities in the region: Rufisque (7 %), Guédiawaye (11 %) and Pikine (16 %). The other regions of Senegal and foreign nationals represent 13% and 2% of the notifications, respectively. Part of the deaths (4 %) could not be classified due to the lack of information on the deceased's place of residence.

Table 2.8: Place of residence in the Main Civil Registry Centre (2016-2017)

Identification of places of residence	Numbers	Percent (%)
The city of Dakar neighbourhoods	1,217	47
Other neighbourhoods of the region of Dakar	875	34
Other regions of Senegal	352	13
Other countries	57	2
Hard to classify neighbourhoods and communes	54	2

Neither addresses nor informed communes	55	2
Total	**2,610**	**100**
Source : Main Civil Registry Centre of the city of Dakar (2016-2017)		

This analysis concerns only the 1,217 deaths recorded in the recruitment area of the MCRC in Dakar. They have been compared to the number of inhabitants in order to estimate the apparent crude mortality rate by neighbourhood. The map of the spatial distribution of the number of deaths by neighbourhood shows disparities at the city scale, without any apparent effect of the physical proximity to the civil registration centre (Fig. 7 and Fig. 8). The spatial mortality quotients reveal the same trends in the distribution of the apparent death rates per thousand inhabitants with significant inequalities in the city (Fig. 9 and Fig. 10). The Moran's test reveals absence of spatial autocorrelation, with observed value (-0.004) lower than expected value (-0.002), and associated p.value of 0,5. This result indicates that the spatial distribution of mortality is not clustered. Resolving registered deaths at neighbourhood level reduced the number of observations for the Moran's test. Furthermore, there is an important number of neighbourhoods without any recorded death, which raises the issue of overdispersion. The small number of observations does not allow sufficiently robust spatial statistical analyses. Such spatial analyses must also assume that data completeness is spatially homogeneous, which is most likely not the case.

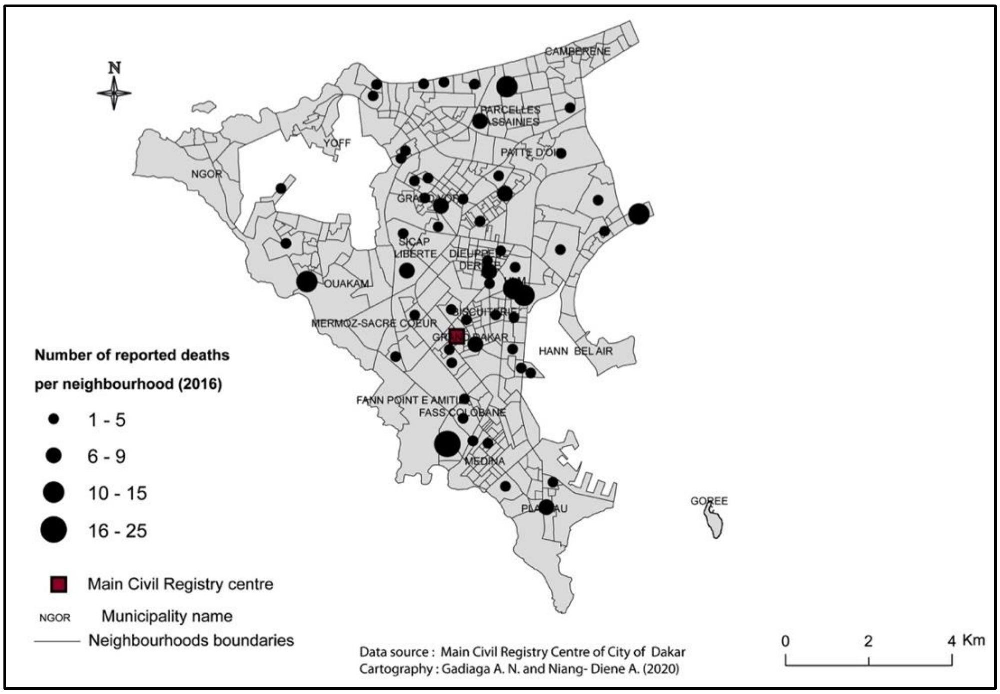


Figure 2.8: Number of reported deaths by neighbourhood in Main Civil Registry (2016)

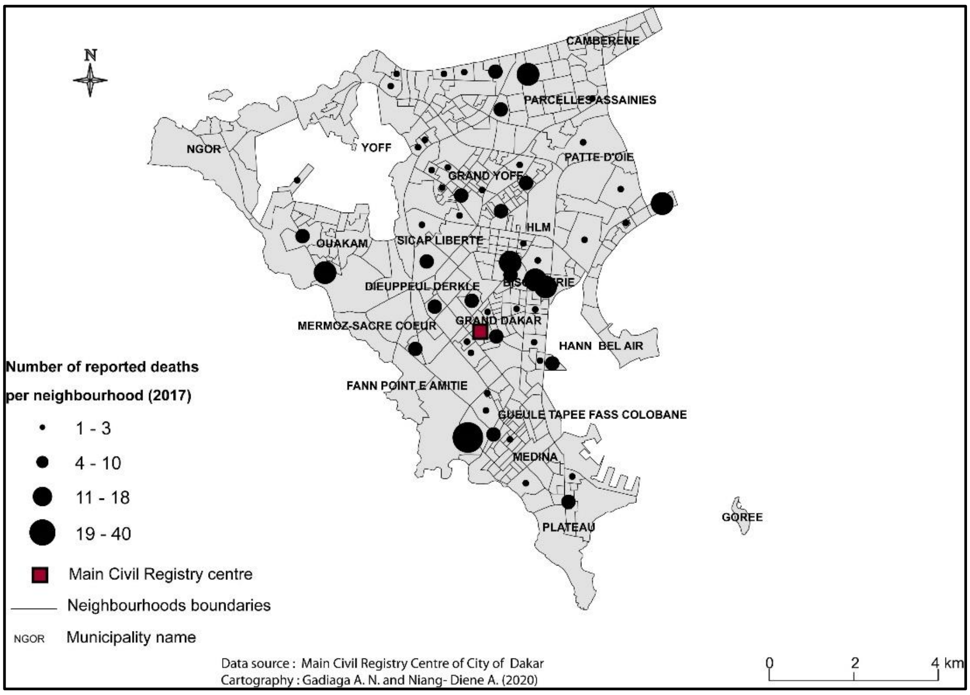


Figure 2.9: Number of reported death by neighbourhood in Main Civil Registry (2017)

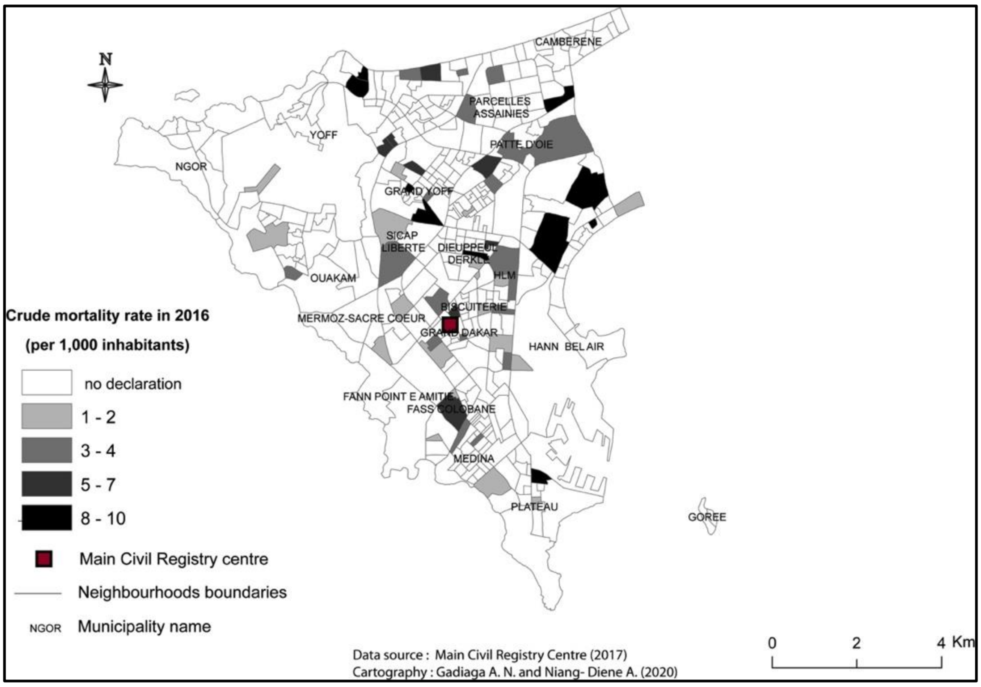


Figure 2.10:Crude mortality rate by neighbourhood in 2016 (per 1,000 inhabitants)

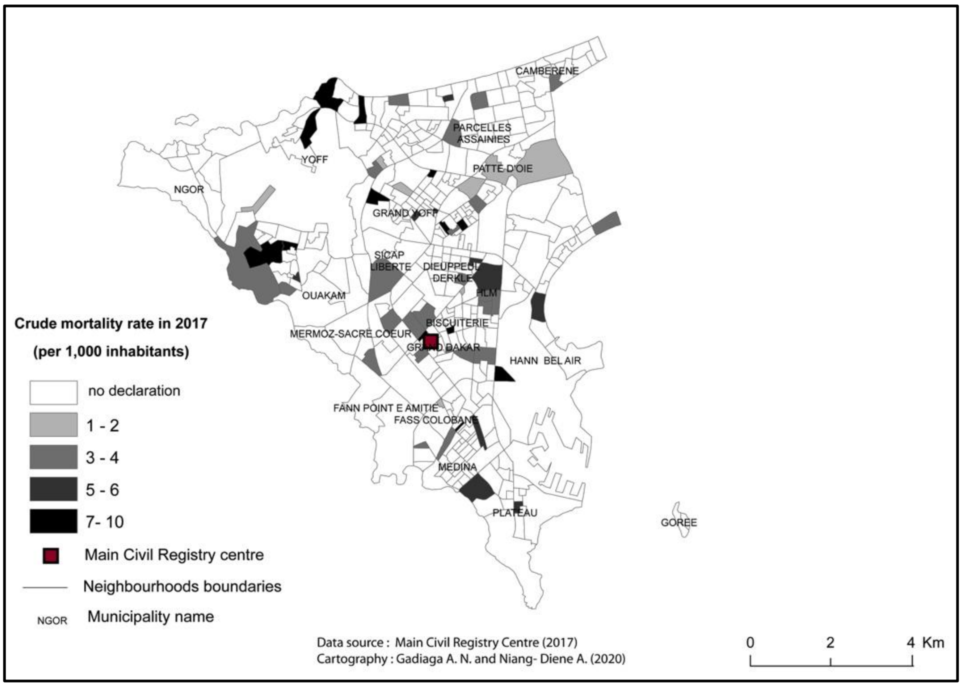


Figure 2.11: Crude mortality rate per neighbourhood in 2017 (per 1,000 inhabitants)

2.4.3.2. Reported places of death

We attempted to analyse places of death from both MCRC and DHIS 2 databases. Aggregated data highlight important limitations in recording this information, particularly in 2016. Only 6% and 43% of places of death were recorded at the civil registry level in 2016 and 2017 respectively (Table 2.9). It is surprising that the largest number of places of death identified are the Le Dantec (11%) and Fann Hospital (7%), which have their own secondary civil registry centres. In Senegal, one strategy in improving declaration of births and deaths at civil registration was to implement civil registration offices in the hospitals. Secondary civil registry centres have been set up in the main hospitals in the Dakar region.

Table 2.9: Places of death reported to the Main Civil Registry Centre (2016-2017)

Place of death	**2016**		**2017**	
	Numbers	Percent (%)	Numbers	Percent (%)
Undetermined	1138	94	753	57
University Hospital Le Dantec	17	1	165	11
University Hospital of Fann	15	1	101	7
Gaspard Kamara Health centre	22	2	96	6
Child health hospital Albert Royer	4	0,3	79	5
Principal Hospital	6	0,5	68	5
Abass Ndao Hospital	0	0,0	21	1
At home	1	0,1	14	1
Private Clinics	5	0,4	65	4
Others public health centres	5	0,4	24	2
Grand Yoff Hospital	1	0,1	9	1
Total	**1214**	**100**	**1489**	**100**
Source : Main Civil Registry Centre of the city of Dakar (2016-2017)				

The mapping of death volumes by health care structure in DHIS 2 shows that these two hospitals did not report any mortality in 2016 and 2017 (Fig 15 and 16).

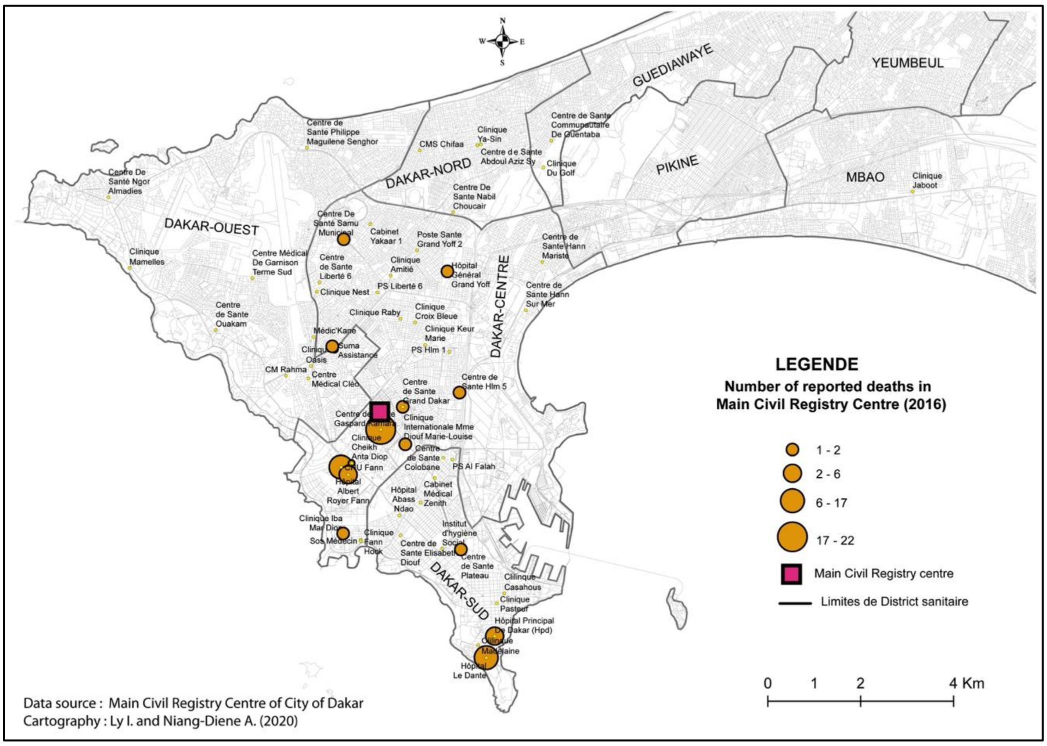


Figure 2.12: Number recorded deaths in Main Civil Registry Centre (2016)

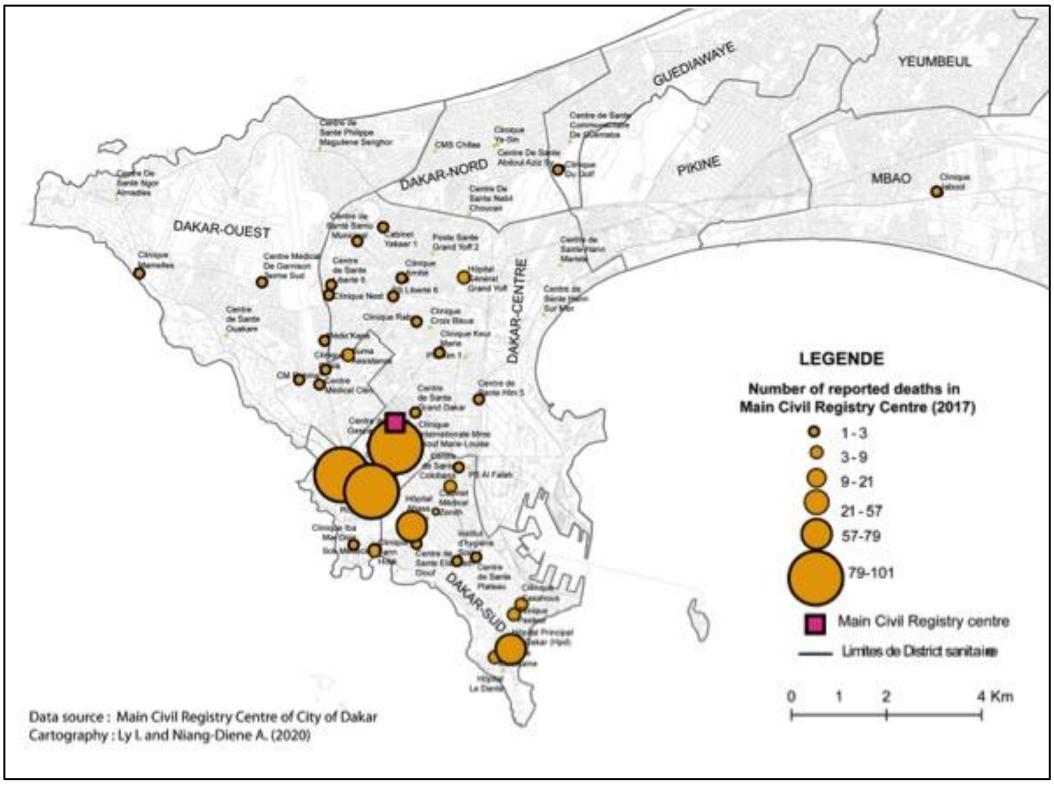


Figure 2.13: Number recorded deaths in Main Civil Registry Centre (2017)

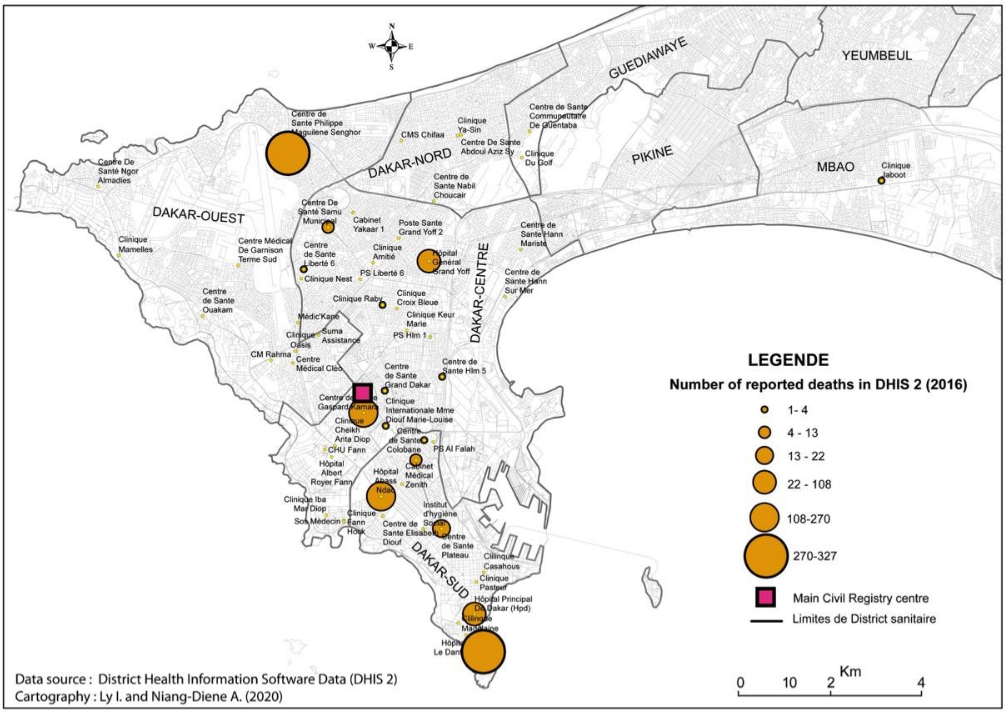


Figure 2.14: Number recorded deaths in health care facilities in DHIS 2 (2016)

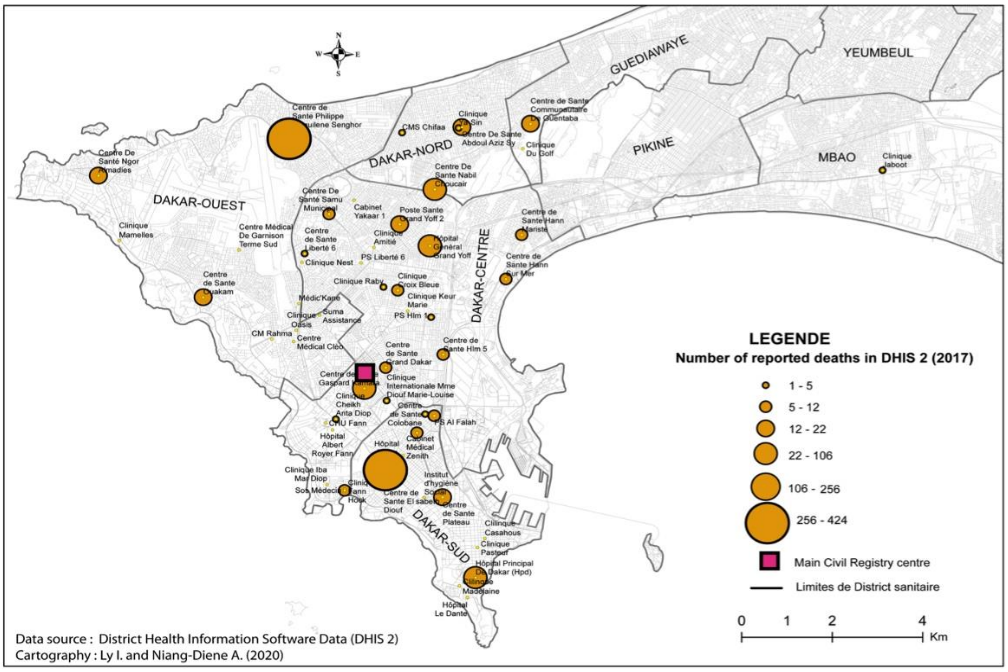


Figure 2.15: Number recorded deaths in health care facilities in DHIS 2 (2016)

The distribution of the volumes of perinatal mortality (stillbirth and 0-28 days death) that are better documented in DHIS 2 indicates that the most important counts, outside of the Abass Ndao hospital, come from the reference health centres of the health districts Dakar-ouest (Philippe Maguilene Senghor), Dakar-centre (Gaspard Kamara) and Dakar-nord (Nabil Choucair). These four health structures concentrate a total of 72% of perinatal deaths (Table 2.10).

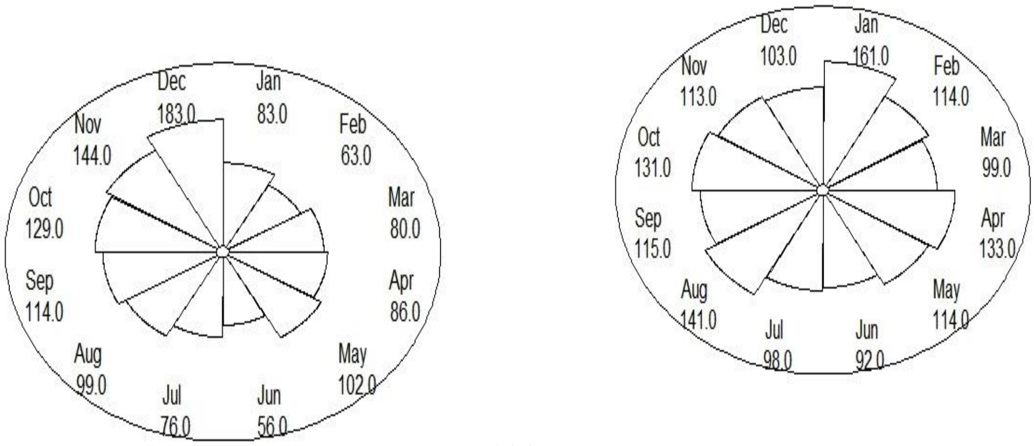
Table 2.10: Number of perinatal deaths recorded in health care facilities of the city of Dakar in District Health information Software Data (DHIS 2)

Health facilities	2016		2017	
	Numbers	Percent (%)	Numbers	Percent (%)
Abass Ndao Hospital	327	25	424	28
Philipe M. Senghor Health centre	270	20	256	17
Gaspard Kamara Health centre	223	17	216	14
Nabil Choucair Health centre	139	10	181	12
Others public health centres	149	11	166	11
Grand Yoff Hospital	108	8	120	8
Principal Hospital	72	5	106	7
Private Clinics	40	3	39	3
Total	1328	100	1508	100
Source : District Health Information Software 2 (2016-2017)				

On the other hand, no deaths have been recorded for the Dakar-sud district, whose reference centre is the Institute of Social Hygiene. The other public establishments are health posts and secondary health centres (11 per cent) and two national hospitals (Grand Yoff with 5 per cent and Principal with 8 per cent). Only 3 per cent of mortality comes from the private sector (private doctors' surgeries and clinics).

2.4.4. Temporal variations in mortality

The study of seasonal variations of deaths can help identify periods of the year with higher risk of death. Aggregated data on monthly recorded deaths at civil registration were analysed for 2016 and 2017. Analysis of the seasonal distribution of death volumes using data from the MCRC shows differences between 2016 and 2017. During 2016, the largest numbers of registrations were made from September to December and then in May (Fig.11).



2016 2017

Figure 2.16: Seasonal variations in the number of deaths declared at the Main Civil Registry Centre

Source: Main Civil Registry Centre of the city of Dakar (2016-2017)

Slightly more than half of the deaths (55%) were counted during these five months. For the year 2017, the amplitudes are less significant. Peak mortality is reached in January followed by August and April. It should be noted that, whatever the year, the lowest records are made during the months of March, June and July. There are questions about the January-February to August differences between the two years. Several studies on morbidity and mortality in Senegal and Dakar actually show a greater use of care for acute respiratory ailments, diarrhoeal diseases and malaria during these two seasons (Salem G. 1998; Dramé, FM, 2006 a; Sy I. 2006; Niang-Diene A., 2019).

2.4.5. Causes of deaths

Apart from the high proportions of stillbirths, the ranking of death causes is very different between the civil registration and the health care system (Fig. 12). In the civil registry data, almost seven out of ten deaths are unexplained. « Death from illness » is mentioned for 37.8% of deaths and 31% have no reason. « Other causes », which accounts for 26% of the information in the health care system, should also be considered (Fig. 13). Such imprecise data make it impossible to assess the risk factors associated with mortality in the city of Dakar. Nevertheless, they do draw attention to the burden of diarrhoeal, respiratory and cardiovascular diseases as well as perinatal mortality.

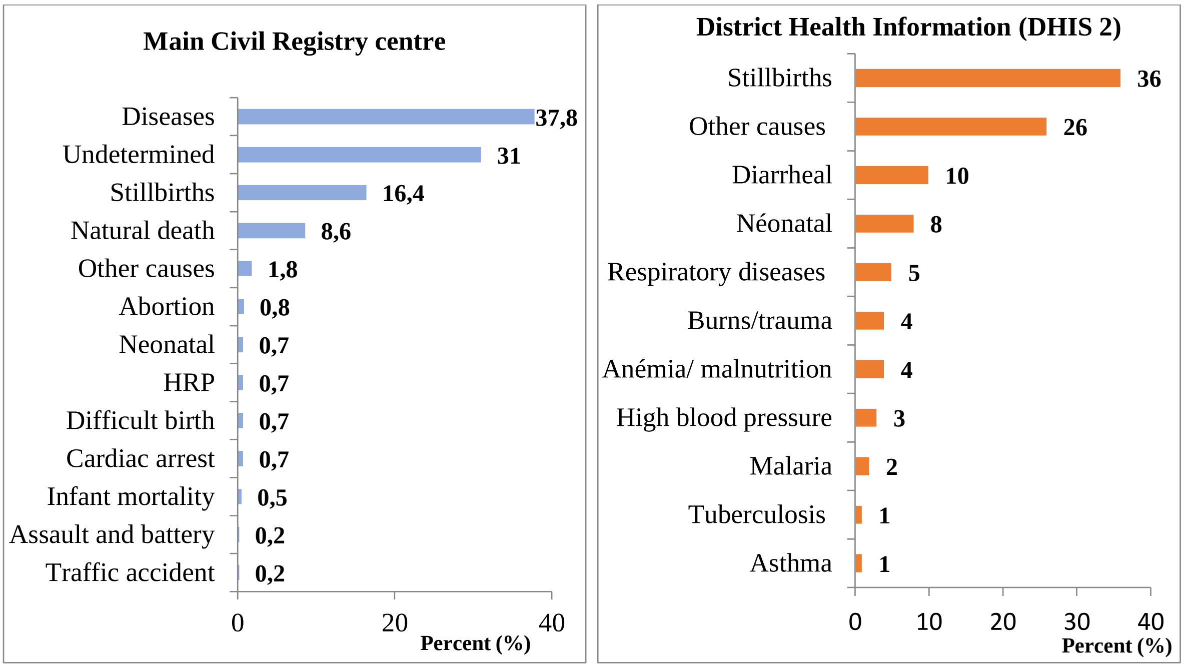


Figure 2.17: Causes of death by source (2016-2017)

Analysis of the causes of neonatal (0-11 months) and infant (12-59 months) mortality shows similar trends (Table 2.11). Overall, more than 70% of the recorded causes are not associated with any particular morbidity. However, the data show that the health care system records a large proportion of child deaths attributable to diarrhoeal diseases (44.7 % for 0-11 months and 25.2 % for 12-59 months), malnutrition (5.1%and 13.2%) and respiratory diseases (14.4 % and 11.2 %).

Table 2.11: Causes of neonatal and infant deaths (%)

Main Civil Registry Centre (MCRC)			District Health information (DHIS 2)		
Causes of death	0-11 months	12-59 months	Causes of death	0-11 months	12-59 months
Childbirth	3,6		Acute surgical abdomen	2,5	

Cardiac arrest	0,8	2,0	Access severe malaria		0,2
Others	2,4	6,1	Anemia		2,7
Abortion	3,2		Asthma	0,5	1,2
Respiratory distress	0,4		Other trauma	1,5	4,4
HRP	3,6		Burns	1	2
Intoxication	0,0	1	Diarrhoea	44,7	25,2
Diseases	44,8	71,4	Low respiratory disease	14,4	11,2
Malnutrition	0,4		Malnutrition	5,1	13,2
Natural death	9,7	4,1	Other causes	30,3	39,9
Infant mortality	1,6				
Undetermined	29,4	15,4			
Sources: Main Civil Registry Centre of the city of Dakar (2016-2017) District Health Information Software 2 (2016-2017)					

2.5 Discussion

The main objective of the present research was to evaluate the quality of the two main sources of mortality data: civil registration data and the new health information system of the Ministry of Health (DHIS 2). More specifically, our aim was to assess to what extend spatial and temporal analyses can be performed based on these existing mortality data.

The study and comparison of mortality data extracted from the civil registry database and the healthcare system revealed important limits that drastically constrain spatial and temporal analyses. In Senegal, only 30.8% of deaths were recorded in the civil registration, according to the results of the 2013 census with real disparities depending on the region, place of residence and age group (RS/ANSD 2016). The region of Dakar appears to be privileged with only 11.4% of cases of death not declared, the same source points out. As stated by Salem (1998), these measures are strongly influenced by the quality of the recording but provide possibilities for evaluating mortality rates on scales that take into account the very great heterogeneity of the city.

This work made it possible to compare data from the health information system and the civil registry in a context where several strategies have been implemented to improve these two mortality registration systems. It provides an initial assessment of their completeness and quality in terms of volumes, spatial coverage and parameters for characterizing mortality (age, gender, periodicity and causes). These preliminary results on the comparison of aggregated data show important omissions in these two sources that make it difficult to estimate robust mortality quotients. They also present deficiencies for a spatialized analysis of mortality at the scale of the city of Dakar. One of its main weaknesses is the non-integration of statistics from secondary centres into the main centre database. Secondly, the statistics from the available secondary centres are not disaggregated. They can only be used to estimate the volumes of deaths recorded per year.

The DHIS 2 data does not allow a spatial location of deaths because the platform does not provide the residential address of deceased persons, even at the neighbourhood level. To obtain this information, it would be necessary to go back to the daily registers of the care structures.

It is not possible to estimate the number of deaths that were double recorded in the two systems. There are also biases relating to toponymy, homonymy, data entry errors and the individual behaviour of populations in these particular circumstances.

Another issue for spatial analyses is the fact that deaths are not always registered in their reference centre i.e. the secondary civil registration office. Analysis of MCRC's data revealed that more than half of addresses of deceased persons are located outside of the city of Dakar. This led to incomplete spatial coverage of civil registration database in secondary centres. In addition, the spatialization of the data in the recruitment area of the main centre shows disparities that can be explained by the quality of data recording but also by other factors that deserve further study.

The two data sources highlight high perinatal and infant death volumes in the city of Dakar. In the 2017 continuous DHS, 226 stillbirths and 236 early neonatal deaths were recorded in the five years prior to the survey. The perinatal mortality rate is therefore estimated at 41 ‰ pregnancies of at least seven months (RS/ANSD 2017). The large discrepancies noted in the stillbirth and infant mortality probabilities between the civil registry and the healthcare system, as well as in the spatial distribution, provide elements for discussion. The deficiencies in the apparent mortality volumes can be explained by several non-exclusive factors, including the quality of the data recording (RS/ANSD 2015; Hane 2017; Furtado 2018), the activity of the healthcare system and the recourse behaviour of the populations. As noted by Woods (2008), account must be taken of the international variability in definitions of stillbirths and the difficulties in the methods of estimating these late fetal deaths, which represent a significant share of mortality in all populations, whose life expectancy is low. The World Organization's report on threshold approaches to fetal viability reflects clinical standards in the assessment of stillbirths and early neonatal deaths (WHO 2006). The differences obtained in this research can therefore be considered an illustration of these medical and methodological differences.

It has been impossible to analyse the data from the secondary civil registration centres because of the lack of completeness. The data that could be spatialized represent only 17% of the statistics collected in the civil registry centres. The lack of spatial autocorrelation in the geographical analysis of civil registration deaths is not in line with our expectations. The above-mentioned small number of spatialized observations and the suspected spatial heterogeneity in data completeness does not allow any robust spatial autocorrelation test to be performed. In addition, there was a large number of neighbourhoods with no reported deaths. This raises the problem of overdispersion, a common problem in epidemiological studies. Spatial epidemiological models are particularly prone to spatial autocorrelation (Linard and al. 2007), because of neighbouring effects between spatial entities but also because these spatial entities, resolved to finer scales, are subject to large variations (Linard and Tatem 2012). Thus, although civil registry data are more appropriate for a geographic approach to mortality, as opposed to DHIS data, efforts are needed to improve civil registration data before using them for that purpose. There is a great need for a system reorganisation with better harmonisation between the central level and the decentralised services.

These methodological limitations as well as problems of comparability and exhaustiveness are obviously at the heart of the problem of this article. They merit further investigation into the activity of civil registry centres, data recording methods and the representativeness of users. They also challenge decision-makers on the need for further research on the determinants of newborn and child health.

2.6 Conclusion and outlook

This article is a first attempt at analyzing the quality of mortality data collected by the Civil Registry and DHIS 2 at the intra-urban scale. So far, the most widely used mortality statistics were those from major national surveys (demographic and health surveys and population censuses) and health and demographic surveillance system. The information provided by these surveys rarely goes beyond the regional level and does not allow for finer scale analyses.

Our results revealed an under-recording of deaths in both the Main Civil Registration Centre and the DHIS 2. These methodological and source quality problems make a spatial analysis of these data on the scale of the city of Dakar difficult. These findings remind us of the limitations of these data for the evaluation of mortality and the construction of indicators that can help improve epidemiological surveillance. They concern only the population that has been able to access the health care system and civil registration. In spite of these shortcomings, the work carried out has nonetheless enabled an exploration of the spatial inequalities in the risks of death in the recruitment areas of the health care structures and the MCRC. Although these data are not sufficient to determine public health priorities, they still provide information on mortality patterns. Mortality volumes show the high burden of stillbirths, neonatal and infant deaths. They index late recourse to maternal and infant care or failures in the care of parturients and newborns.

However, since the causes of death are not systematically specified, they are still unknown for at least 70 per cent of the deceased. Nevertheless, they are still useful indications of mortality risks due to diarrhoeal, respiratory and cardiovascular diseases. It is therefore necessary to pay greater attention to the quality of statistics and methodological problems in the collection of health information. In this sense, civil registry data can help to improve indicators on the morbidity profiles of the population. Even if extrapolation is not possible with these data, they can make a significant contribution to other research on intra-urban health inequalities.

Civil Registry and Health Information System databases are the only ones that can be used for continuous observation. They can provide decision-makers with better knowledge of mortality and morbidity patterns to guide health actions. The results of this study encourage further investigation to better highlight the social-spatial factors in differential mortality. In this context, the tools of geography can contribute to a better analysis of the interactions between territorial, epidemiological and socio-spatial dynamics. However, for this to happen, the procedures for recording and processing data both at the civil registry level and in the DHIS 2 would need to be improved.

Chapter 3

Neighbourhood-level housing quality indices for health assessment in dakar, senegal[5]

[5] **Adapted from :** Gadiaga, A. N., De Longueville, F., Georganos, S., Grippa, T.; Dujardin, S., Diène, A. N., Masquelier, B., Diallo, M., & Linard, C. (2021). ***Neighbourhood-level housing quality indices for health assessment in Dakar, Senegal, Geospatial Health,*** 16(1) https://doi.org/10.4081/gh.2021.910

In the city of Dakar, accurate and detailed spatial data have been gathered from very-high resolution (VHR) satellite imageries. Spatial heterogeneity of the urban area has been highlighted by these land cover-derived VHR satellite imageries, as well as variation in environmental and socioeconomic conditions, with a mapping of residential land use areas. Furthermore, reliable census data which provides socioeconomic indicators and allows direct estimates of mortality rate are available for Dakar. Based on these data, this chapter carries out a typology of the Dakar's neighborhoods and analyse association between overall crude death rate and neighborhood housing quality profiles

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 (PCA) 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.

KEYWORDS *Census data, Land cover/land use data, Mortality rate, Neighbourhood, Dakar.*

3.1 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 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 1998b). 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 analyze 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, however, 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 (Adjei and Kyei 2013; Bawah and Zuberi 2004.; Suglia et al., 2011), on the one hand, and remote sensing data (C. N. Thomson et 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 (Gething et al., 2015; Tapiador et al. 2011; 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 satellite imagery of very high resolution (VHR). 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 therefore 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).

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.

3.2 Material and methods

3.2.1 Study area

The Dakar region is divided in four departments: the city-centre in its southern part (Dakar), Pikine, Guediawaye and Rufisque (Figure 3.1). 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 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, 2014). 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.

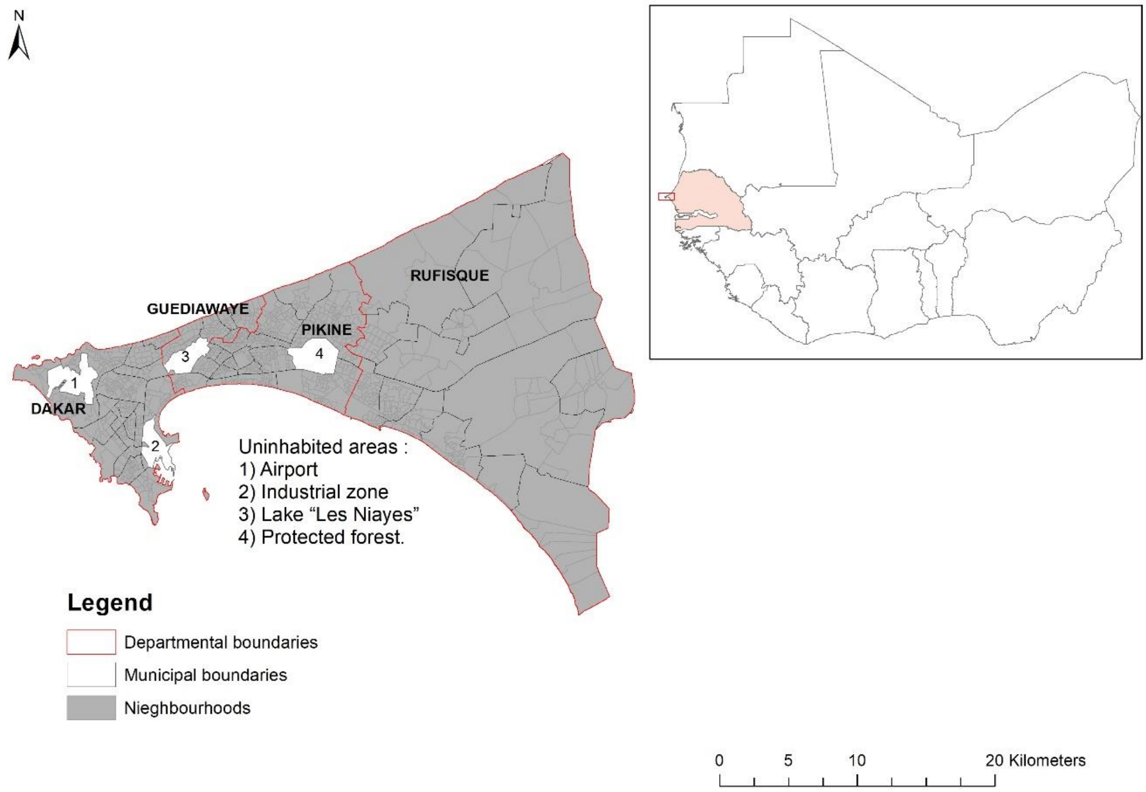


Figure 3.1: Geographical location of Dakar and its administrative units.

3.2.2 Data

3.2.2.1 Census data

Census data for the year 2013, recorded by enumeration area (N= 4,189), 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=1,347), 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.

3.2.2.2 Remotely sensed data

Open access land cover (LC) and land use (LU) maps derived from VHR satellite data for the year 2015 were used in this study (Grippa et al., 2018a). Local segmentation and random forest classification according to Grippa et al. (2017b) allowed the identification of 11 LC classes (Figure 3.2a). 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 3.2b). 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.

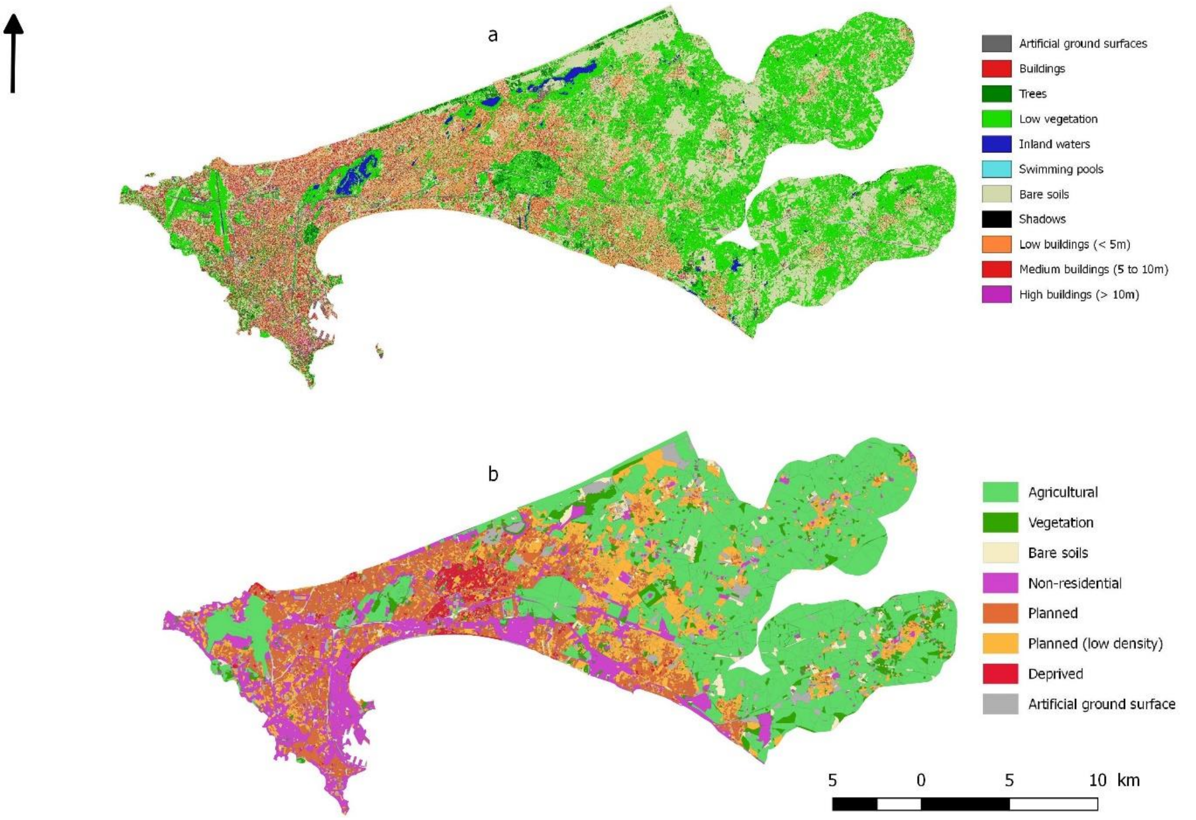


Figure 3.2: Land cover (a), and land use (b) maps of Dakar (Data source: https://zenodo.org/record/1290800#.W7xWQ_a2w2w)

3.2.3 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 et Zuberi (2004), in combination with Dakar's detailed LC and LU data (Grippa et al. 2018). Table 3.1 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 the type of floor covering high-quality and poor-quality) were combined. They were selected because (i) the coefficient of variation were 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 with 0.5 m spatial resolution were 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 1998b).

Table 3.1: 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

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 (Bawah and Zuberi, 2004; Lo and Faber 1997; Li and Weng 2007). Using the FactoMineR package of the R statistical software (Lê, Josse, et Husson 2008), PCA was applied to the 16 variables listed in Table 3.1. 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=1,347). 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.

3.2.4 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.

3.3 Results

3.3.1 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 3.2). 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 3.2.

Table 3.2: Correlation coefficients between input variables and four components extracted from PCA

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

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.

Variables with noteworthy positive loadings on Component 2 were "agricultural vegetation", "natural vegetation" and "non-residential build-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 a "poor-quality property".

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 area".

Figure 3.3 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 3.3a). 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 3.3b). In the centre, component 2 also highlights pockets of vegetation around the protected forest and the wetlands, such as the lake "Les Niayes" (Figure 3.1) and along the western coast. Component 3 represents the poor-quality properties and higher scores are located in the rural areas of Rufisque (Figure 3.3c). 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 3.3d 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.

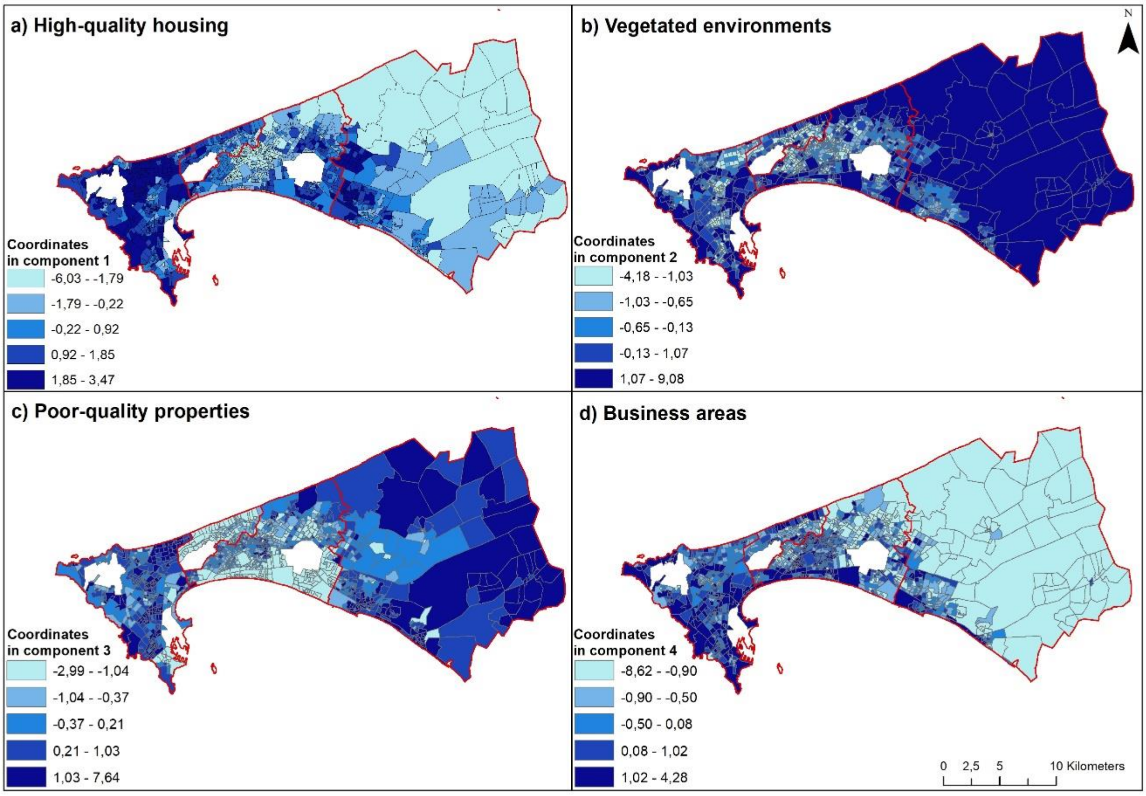


Figure 3.3: Spatial distribution of the four components defined by the PCA

(a) high-quality housing (Component 1); (b) vegetated environments (Component 2); (c) poor-quality properties (Component 3); (d) business areas (Component 4). Discretization in quantiles.

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 3.3 presents v-test results and Table 3.4 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 where 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 3.4).

Table 3.3: V-test statistic of components by cluster

PCA component / Cluster	**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

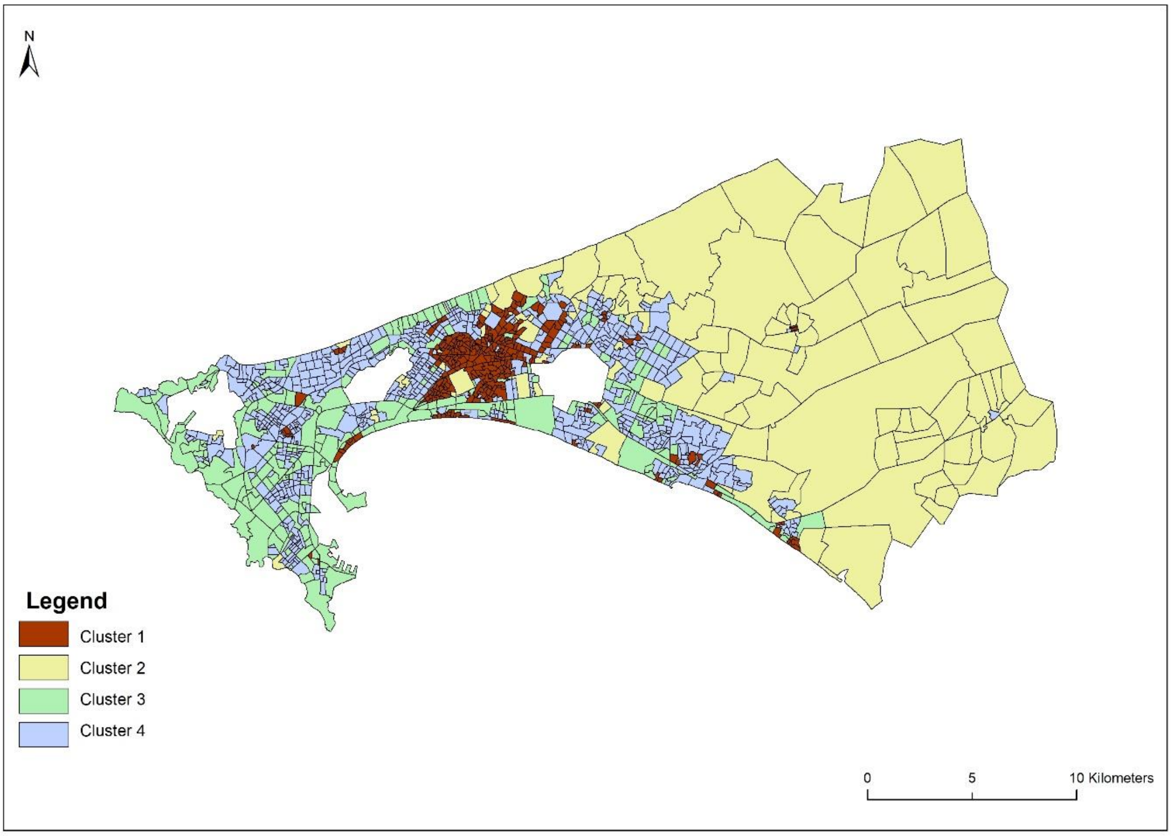


Figure 3.4: Distribution of neighbourhoods by clusters

Cluster 2 was positively associated with PCA-component 2, i.e. vegetated environments. It had the largest proportion of homeowners (40% of households) and the fewest number of neighbourhoods (116 out of 1325), mainly located in the eastern and less densely populated areas. Table 3.4 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 1,325). Clusters 3 and 4 spread outside the city centre, in the department of Rufisque.

Table 3.4: 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 of 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.

3.3.2 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 3.5). The Kruskal-Wallis test showed significant differences between clusters (p-value <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).

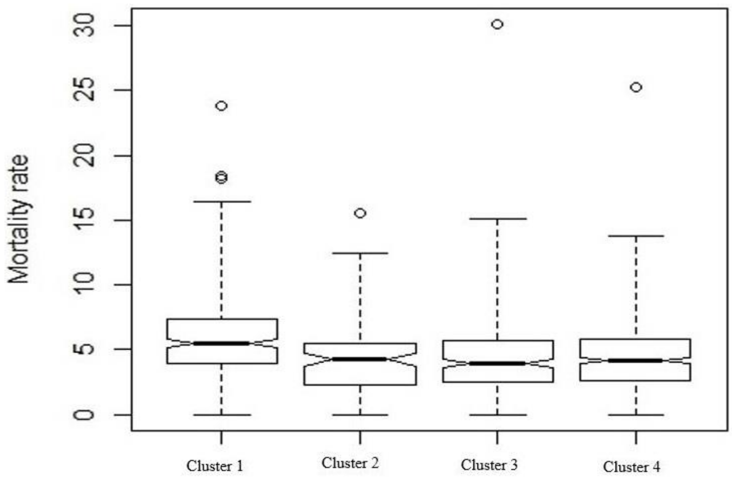


Figure 3.5: Distribution of crude mortality rates by cluster

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. value < 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-value = 0.005). The relationships between the mortality and the remaining components were not statistically significant (p-value = 0.79 for the correlation between the mortality rate and component 3 ; p-value = 0.13 for the correlation between the mortality rate and component 4).

3.4 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. It provides an original typology of neighbourhoods based on fine-scale housing quality maps of Dakar. 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, as 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 was carried out for the four departments of the Dakar region and at the neighbourhood level, which is the most relevant level for health interventions. We also analysed 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. 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), and 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, these cannot be computed at the neighbourhood level from data collected in the census due to the small number of deaths in the 12-month reference period. 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.

3.5 Conclusion

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 images 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.

ACKNOWLEDGMENTS

This research was funded by ARES-CCD (https://www.ares-ac.be) in the frame of the ASSESS project (http://assess-sn.org/). The authors gratefully thank the MAUPP project (https://maupp.ulb.ac.be/) and Senegal census bureau (National Agency for Statistics and Demography) for the access to the data.

Chapter 4

Determinants of geographic distribution of age-specific mortality in Dakar, Senegal[6]

[6] Adapted from: Gadiaga, A. N., Georganos, S., De Longueville, F., Salem, G., Diène, A. N., Fall, A., Masquelier, B., Diallo, M., & Linard, C. **Determinants of geographic distribution of age-specific mortality in Dakar, Senegal. (In preparation)**

In the previous chapter, we analysed the impact of living conditions on health by examining the relationship between neighbourhood housing quality profiles and crude mortality rate. Statistical distribution analysis and hypothesis tests were used.

In this chapter we look more into the relationship between living conditions and health, with a focus on age-specific crude mortality rate. A set of variables related to neighbourhood contextual characteristics is used, and geographically weighted regression model is applied in order to capture the varying relationships between infantile, adolescent, adult and old age mortality, and neighbourhood contextual characteristics.

4.1 Introduction

Mortality is an objective and measurable indicator for population health status. A low mortality rate, especially in the lowest age groups, reflects the high socioeconomic development level of the population considered (Cutler et al. 2006) as well as improvements in sectors such as hygiene, sanitation, and access to healthcare services. On the contrary, countries facing burden of infectious and chronic diseases, poverty, and worse socioeconomic and environmental conditions are characterized by high mortality rates, especially in lowest age groups. For that reason, international measures of well-being, namely the Human Development Index and the Sustainable Development Goals incorporate life expectancy and mortality rates as important indicators of well-being (WHO 2016). Despite a decline in mortality rates these recent years, Sub-Saharan Africa (SSA) is the region of the world with the highest level of mortality for all age groups except the elderly (World Population Prospects - Population Division - United Nations, 2019). SSA is also subject to important spatial variations of mortality rates and associated socioeconomic (Quentin et al. 2014) and environmental risk factors (Dos Santos et al. 2015), especially within cities. The analyses of age-specific mortality rates provide a more detailed description of mortality patterns than overall mortality rates (Wang et al. 2017), especially as leading causes of death – and hence determinants of mortality – vary considerably if we decompose mortality rate by age-specific groups (El Bcheraoui et al. 2020). Different age-groups lead to characteristic determinants of mortality: for the lowest age-groups, the impact of environmental contamination is preeminent, as the lowest age-groups are heavily burdened by infectious and parasitic diseases; for adults and elderly, lifestyle factors and behaviours are more important than environmental contamination, as adult and elderly health are more affected by chronic diseases.

Notwithstanding the vast majority of studies on potential drivers of mortality in urban environments in SSA, three key aspects remain challenging for a detailed assessment of mortality variations at intra-urban level. First, the sampling strategy applied in commonly used survey data is not adapted. Several authors have evidenced that the methodology designed in the Demographic and Health Surveys (DHS) in low- and middle-income countries -the largest sample data used for health studies- is flawed, that the sample is often too small for accurate estimates of at-risk groups (Madise et al. 2012; Quentin et al. 2014). Furthermore, the poorest urban dwellers are left out from the sample due to the statistical or demographic weighting scheme used (Günther et Harttgen 2012; Thomson et al. 2021), whilst the largest mortality differentials are found in the living areas of these poorest and vulnerable urban dwellers. Second, infant (0-1 year) mortality has gained much attention in the literature, but little is known on mortality in other age-groups. This shortcoming is related to the absence of good quality vital statistics and census data, which often prevent direct estimates of age-specific mortality rates. Indirect methods based on available child mortality data are then generally used to estimate infant mortality (Wang et al. 2017; Clark 2019). This is based on the assumption that child mortality is usually closely related to adult mortality and inequalities in child mortality are sufficient proxies for understanding adult mortality (Günther and Harttgen 2012). However, the correlation is not this straightforward, as it may be influenced by epidemics that affect age-groups differently. The COVID-19 pandemic is a perfect illustration of such exceptions, with a severe death toll during the first wave until recently for elderly and adult age groups. Third, limited empirical evidence is gained from modelling the determinants of mortality without considering spatial effects. Non-spatial regression models assume a constant relationship between mortality and identified risk factors across space, while spatial autocorrelation often occurs and lead to unobserved heterogeneity.

The present study addresses the issue of spatial heterogeneity in the distribution of age-specific mortality rates in Dakar based on census data and model the complex relations between mortality by age groups (0-14, 15-59, 60 and more) and a set of contextual risk factors using a local non-parametric regression known as geographically weighted regression (GWR).

4.2 Data and methods

4.2.1 Analytical framework

We conducted this study at neighbourhood level (N=1347), the finest administrative unit in Dakar (Figure 1.4).

4.2.2 Dependent variable

The 2013 census database is structured into tables, namely the "Individuals", "Deaths", and "Habitat" tables, among others. The table "Deaths" provides information on households' deaths that occurred during the 12 months preceding the census, i.e. from November-December 2012 to November-December 2013. We used the variable "age at death" and aggregated the number of deaths by neighbourhood, according to the three following age-specific groups:

- Child and adolescent mortality: 0-14 years of age
- Adult mortality: 15-59 years of age
- Old age mortality: 60 years and more.

The dependent variables are therefore the number of deaths in the three specific age-groups, and the logarithm of population counts in each age-specific group was used as an offset variable. The offset variable adjusts the regression model to account for the denominator (Barnett et Dobson 2010) and therefore turns death number into mortality rates.

There is a much higher frequency of deaths among the elderly. The mean number of deaths at neighbourhood level is 1.76, 3.91 and 4.61, respectively for child-adolescent, adult and old age population. As age unit increases, mortality also increases.

4.2.3 Covariates

Independent variables include both socioeconomic and environmental variables derived from the 2013's census database. For each neighbourhood, the proportions of households owning their home, built with tiled roof or concrete slab, covered with an improved or unimproved soil, and using a private tap (or mineral water) as source of drinking and sewage system for sanitation, were calculated. We captured literacy rate by the education level of the head of household. We calculated population density by neighbourhood total area and population density by neighbourhood built-up area. This approach is deemed to better capture the potential double effect of population density, at both neighbourhood scale (i.e. living conditions in the neighbourhood) and building scale (i.e. habitat living conditions). We measured poverty levels of the households by the lack of resources that on the one hand prevent access to health care services and on the other hand purchase of foods. All variables used are listed in Table 4.1.

Table 4.1: Description of variable used in this study

Label	**Variable description**
Child-adolescent deaths	Total number of deaths of 0-14 years
Adult deaths	Total number of deaths of 15-59 years
Old age deaths	Total number of deaths of 60 years and over
Owner	Proportion of household heads being owners
Concrete slab	Proportion of households with roofs made of concrete slabs
Tiled roof	Proportion of households with tiled roofs

High-quality floor covering	Proportion of households with floors covered with tiles or carpet
Poor-quality floor covering	Proportion of households with floors covered with sand or clay
Sewage system	Proportion of households connected to the sewage network
Private tap	Proportion of households relying on water from a house tap or mineral water
Education	Proportion of educated head of households
Poverty	Proportion of household who lacked resources for food expenditure.
Lack of care	Proportion of heads of household who lacked resources for health care
Population density	Population per total area
Built-up density	Population per built-up area

4.2.4 Analysis

We first tested the presence and significance of spatial autcorrelation in the spatial distribution of age-specific mortality rates at the neighbourood level using the local Moran's index, which ranges between +1 (similar values clustered together) and -1 (dissimilar values clustered together). A Moran's index of 0 indicates a random spatial pattern. Among the positive and statistically significant spatial outliers, the local Moran's index allows to distinguish between hotspots (i.e. high mortality rates surrounded by high mortality rates) and coldspots (i.e. low mortality rates surrounded by low mortality rates).

Second, we measured the impact of our covariates on mortality and examined the potential spatial variability of these impacts. As death is a rare event, a Poisson regression model is more appropriate than a Gaussian model (Souris 2019). Following Yang et al. (2013) and Georganos et al. (2017), a GWR model examines the spatial non-stationarity and the spatial varying coefficients by taking into account the location of the observations. Thus, a simple Poisson regression formula, formally written as follows:

$y_i \sim \text{Poisson}\ [N_i \exp(\sum_k \beta_k X_{ki})]$

$$\log(y_i) = \log N_i + \beta_0 + \sum_k \beta_k X_{ki} \qquad (1)$$

can be extended to a Geographically Weighted Poisson Regression (GWPR) in the following equation:

$$y_i \sim \text{Poisson}\ [N_i \exp(\sum_k \beta_k (u_i, v_i) X_{ki})] \qquad (2)$$

where y_i represents the value of the dependent variable and N_i the population count of neighbourhood i. N_i is used as an offset term, u_i, v_i refers to the coordinates of the centroid of neighbourhoods, β_k are the coefficients for the independent variables k of neighbourhoods i, with X_{ki} the related set of values of the independent variables.

Focusing on each particular location i, GWPR performs a number of sub-models equivalent to the number of observations. The study area of the sub-models is a target area defined by a weighting scheme in which nearby observations have a non-zero weight. Specifically, these weights are defined by a continuous function of distance, with observations closer to the location of regression point i weighted more heavily. The window over which the local model is applied is the "kernel" and the maximum distance away from the regression location i within this window is the "bandwidth". An advantage of GWR models is their ability to map local parameter coefficients as well as local model performance (pseudo R^2). T-values associated with local coefficients is a key information, giving evidence of non-stationarity and the significance of the local regression coefficients in specific locations. For each GWPR model, the relationship between mortality and contextual factors is either positive or negative, and cannot be simply summarized as a pattern of a specific target location. Because the number of neighbourhoods is too large, it is easy to lose track of the spatial patterns of mortality when looking at neighbourhoods individually. To solve this issue, we followed the approach proposed by Singh and Masquelier (2018), by performing a clustering on the t-values of each GWPR, and then measuring the influence of contextual factors within identified clusters. We looked separately at clusters of the three GWPR models (0-14 years, 15-59 years, 60 years and more), as areas covered by clusters changed significantly between models. We performed a quasi-Poisson regression model for each cluster at the neighbourhood level. It should be noticed that overdispersion of the dependent variable is not treated in the GWPR. So, at this final stage of the analysis, the quasi-Poisson model is used to address the issue of overdispersion. Cluster-specific quasi-Poisson regressions included the full list of covariates presented in Table 4.1 and illustrate how the relationship between mortality and contextual factors vary over the area. Maps of local spatial autocorrelation were generated using the ArcGis 10.3 software. We used the standalone version of the multiscale geographically weighted regression to perform the GWPR and R statistical software for quasi Poisson regression analyses.

4.3 Results

Local spatial autocorrelation revealed concentration of neighborhoods with excess mortality in the peripheral areas, especially for child-adolescent (Figure 4.1A) and adult (Figure 4.1B) mortality rates. Another important spatial pattern of child-adolescent and adult mortality rates is the presence of coldspots alongside hotspots, particularly in Pikine for child-adolescents and Dakar for adults (Figure 4.1 A and 4.B). Some old age mortality hotspots are found in central Dakar and Rufisque (Figure 4.1C).

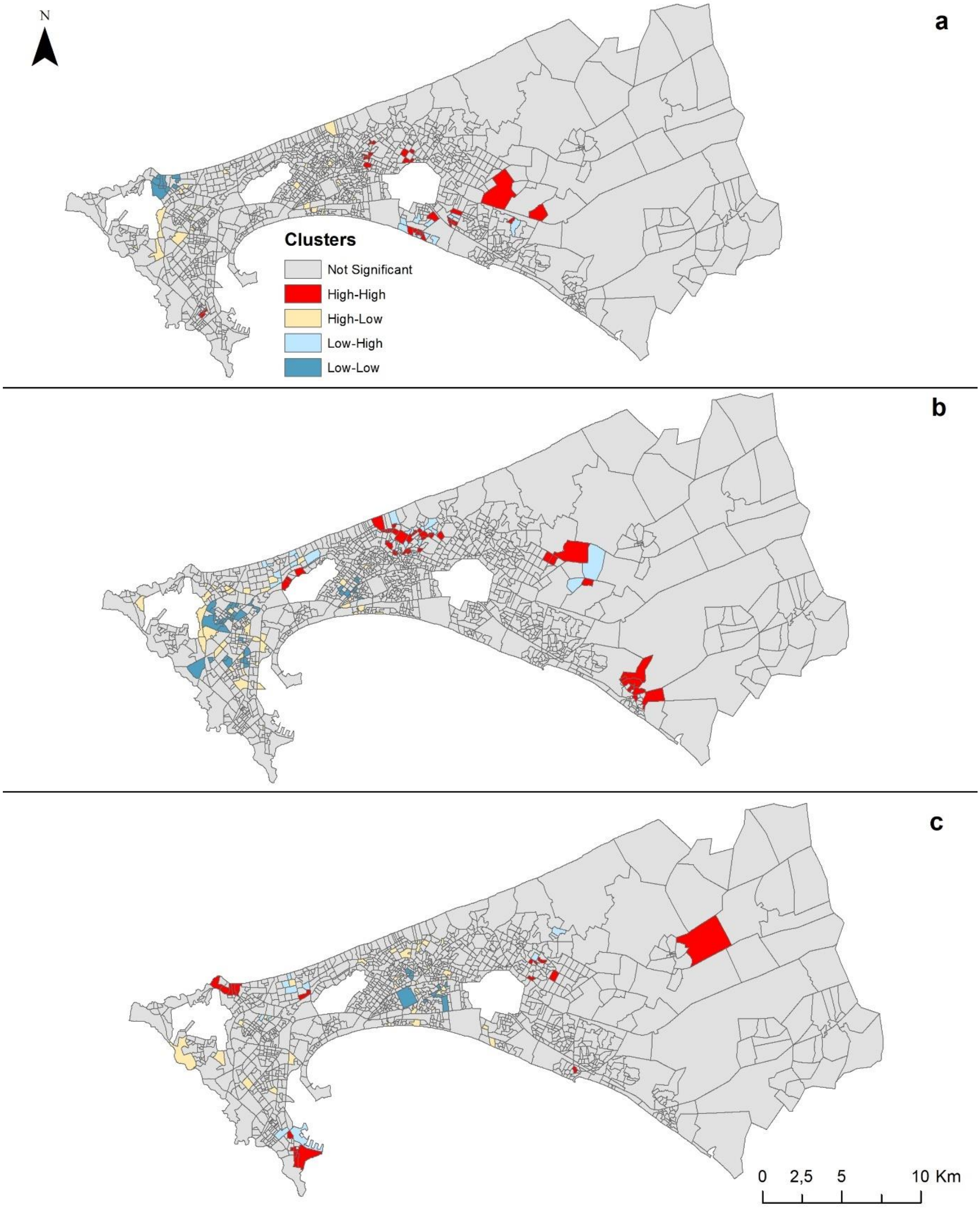


Figure 4.1: Local spatial autocorrelation. a) Child-adolescent mortality; b) Adult mortality; c) Old age

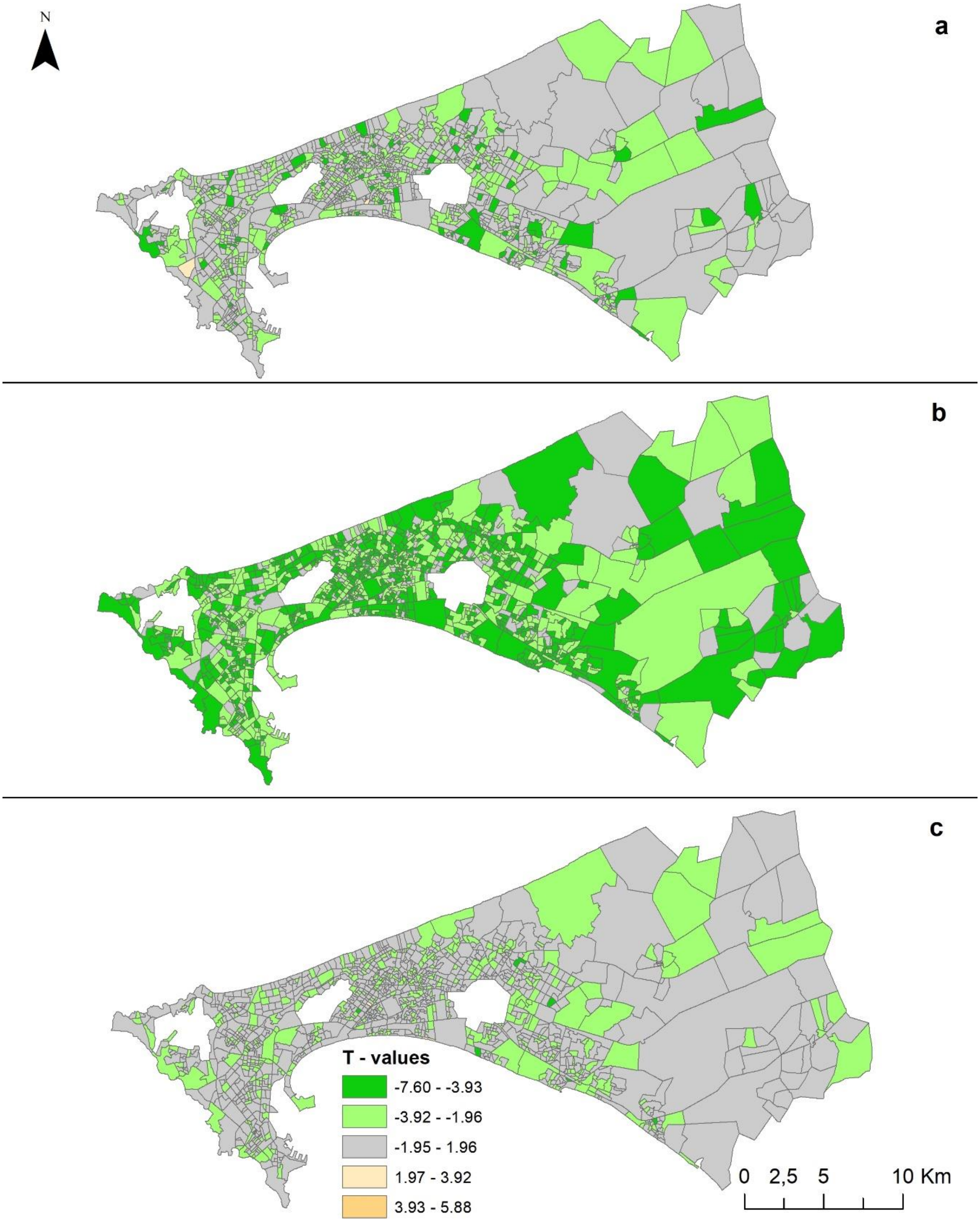


Figure 4.2: T-values associated with local regression coefficients. a) Child-adolescent mortality; b) Adult mortality; c) Old age

The GWPR models suggest that the associations between mortality and the set of contextual variables vary significantly over the study area. Local regression coefficients exhibit important spatial variations, with associated t-values being statistically significant (t = > 1.96 or t < - 1.96) for the vast majority of neighborhoods (Figure 4.2). The R^2 values range between 0.24-0.74, 0.28-0.83, 0.26-0.81, respectively for child-adolescent, adult and old age mortality models.

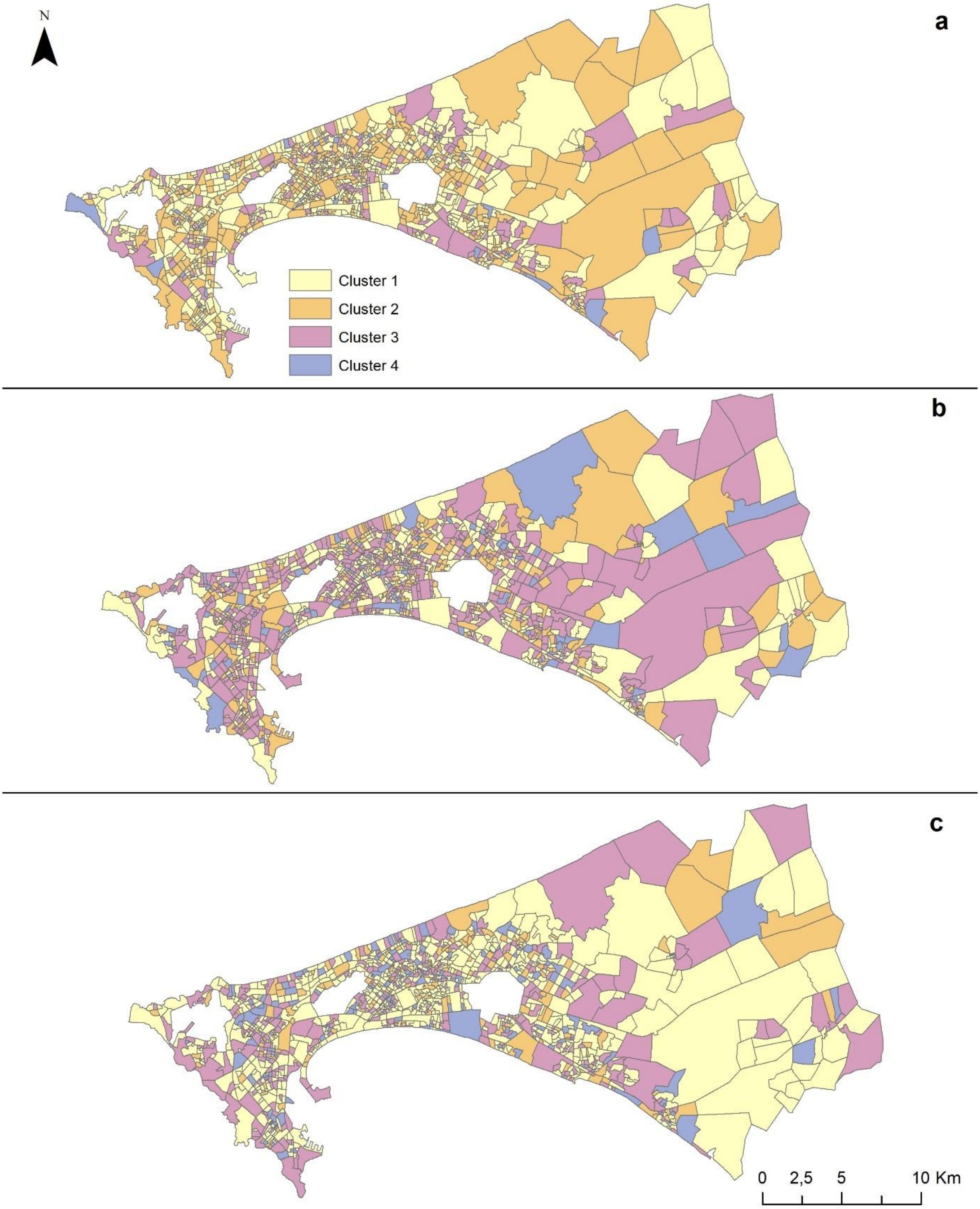


Figure 4.3: Cluster location

Results from cluster-specific quasi-poisson regression models are synthetized in Table 4.2. The influence of contextual characteristics highlights above all the role of population density, which appears to be significantly positively correlated with mortality rates in almost all identified clusters and all age-groups. Other contextual risk factors involved in the spatial variations of mortality are poverty, education, lack of access to healthcare due to lack of income, sewage system, and private tap.

The spatial variability of the relationship between age-specific mortality and contextual risk factors is displayed in Table 4.2, and described in more detail below.

Child and adolescent mortality: Mortality for the 0-14 years is influenced by the variables population density, poverty, education and sanitation. A higher 0-14 years mortality rate in densely populated neighbourhood is the most sizeable spatial pattern that accounts for more than three quarters of all neighbourhoods and that is observed in the downtown, the suburb and the periphery (cluster 1, 2 and 4). In addition to population density, regression analysis in cluster 4 pointed out the influence of education and sanitation, with higher household education levels and improved sanitation being significantly associated with lower mortality rates. Cluster 4 includes a small number of neighbourhoods in the centre of the department of Dakar and the centre of Rufisque, the municipalities of Sam Notaire and Wakhinane in Guediawaye and in the west, the northwest and south of Pikine. Cluster 3 has the highest 0-14 years mortality rates and extends less over the study area. Here, the only significant contextual risk factor is poverty, with a significant positive association with 0-14 years mortality.

Table 4.2: Cluster-specific quasi poisson regression models for child-adolescent, adult and old age mortality

	0-14 years				15-59 years				60 years and +			
Clusters	1	2	3	4	1	2	3	4	1	2	3	4
Poverty			0.0219*		0.0086**			-0.0311**				
Education				-0.0341*		-0.0143*						
Lack of care								0.0268**				
Sanitation				-0.0117**								
Private tap								-0.0081*				-0.0282*
High-quality covering												
Poor-quality covering												
Tiled roof												

Concrete sab												
Owner												
Built-up density												
Population density	0.0013***	0.0010***		0.0023***	0.0011***	0.0016***	0.0016**	0.0014*	0.0009*	0.0011***	0.0008***	
R^2	0.91	0.91	0.95	0.99	0.97	0.98	0.99	0.99	0.96	0.84	0.75	0.99

Adult mortality: Adult mortality is characterised by the variables population density, poverty, education, lack of access to healthcare due to lack of income and private tap. Regression analyses highlighted a significant positive correlation between population density and 15-59 years mortality rates in all 4 clusters. An increase in poverty is also associated with an increase in adult mortality rate in cluster 1, and the higher the education level of the head of the household, the lower the adult mortality rate in cluster 2. Both clusters 1 and 2 are largely dispersed in the downtown and the suburb areas. Cluster 3 only reports the role of population density. Cluster 4 is the cluster with the highest adult mortality rates and, in addition to population density, the output of the quasi Poisson regression analysis emphasises a significant association with the lack of access to healthcare due to a lack of income and safe water sources. In cluster 4, an increase in adult mortality rate is also unexpectedly associated with a decrease in poverty level. This cluster has the smallest number of neighbourhoods and is almost evenly distributed in the departments of Dakar, Pikine and Rufisque. It includes a low number of neighbourhoods from the department of Guediawaye.

Old age mortality: Contextual risk factors associated with old age mortality are population density in cluster 1, 2, and 3 and access to safe water sources in cluster 4. In the three first clusters, there is a significant positive correlation between old age mortality rate and population density and taken together they encompass almost the whole region. The negative correlation between old age mortality rate and access to safe water is the only significant relationship reported by the regression model in cluster 4. Besides, this cluster shows the highest old age mortality rates. The majority of the neighbourhoods of cluster 4 are located in the department of Pikine, in particular in the areas of Keur Massar, Yeumbeul and Mbao, and in a few areas of Dakar and Rufisque departments.

4.4 Discussion

This study is the first in Dakar to investigate the influence of socioeconomic and environmental conditions on age-specific mortality, focusing on mortality in the age-groups 0-14 years, 15-59 years, and 60 years and more. Our findings suggest that determinants of mortality vary according to (1) age-specific groups and (2) where people live. Population density, poverty, education, lack of access to healthcare due to a lack of income, sanitation and private taps are the socioeconomic and environmental drivers of spatial inequalities of age-specific mortality in Dakar. Higher child-adolescent mortality rates are associated with higher population density, poverty, poor sanitation and low education level of the head of the household. These results confirm a study in Ouagadougou, which concluded that the spatial variations in infantile mortality is more complex than a simple opposition between the centre and periphery, and that infantile mortality is greater in neighbourhoods with higher population density and low education level (Soura 2009). Similar effects of education were proven by a study in Kenya, with infants born to parents with a secondary or higher degree of education experiencing a lower risk of death (Liu 2014). This positive effect of education is also highlighted by Pongou et al. (2006) who analyzed risk factors affecting child health in Cameroon, as well as by Bawah and Zuberi (2004) who showed a decreasing effect of household socioeconomic status on child mortality in Botswana, Lesotho, and Zambia. The association we found between higher child-adolescent mortality and poor sanitation confirms the hypothesis that improved household conditions positively impact on child health (Pongou et al. 2006).

The risk factors affecting adult mortality in this study, i.e. population density, poverty and education, are similar to those of child-adolescent mortality. Contrarily, sanitation played no role in determining adult mortality rates. Private tap and lack of access to healthcare due to lack of income became important predictors of adult mortality rates only in certain places. The negative correlation between adult mortality and education is consistent with one of the rare investigations on determinants of adult mortality suggesting that the presence of at least one literate person in a household substantially reduces the death risk of adults (Saikia and Ram 2010).

Our results showed that despite only slight differences between risk factors determining child-adolescent and adult mortality, these become much larger when comparing determinants of lowest with oldest age-groups. Nevertheless, population density was proven to be a primary health risk factor also in the age-groups of 60 and more years.

This study demonstrates that it is relevant to include contextual factors when modelling spatial patterns of mortality. Although neighbourhood characteristics stem from socioeconomic and environmental conditions of a household and not of the dead person himself, these conditions remain representative for understanding the determinants of mortality. Indeed, the health behaviour of a dead person from a household is reflected by the health behaviour of the surviving persons from the same household (Saikia et Ram 2010). Consequently, a higher education level of the head of the household for example, may lead to opportunities to raise income. In addition, a higher income lifts the household out of poverty and improves nutritional intake, leading to a greater resistance to disease. Higher income further enables households to spend more money on healthcare, to enhance living conditions which typically improve housing (safe water source, better sanitation). Lack of access to health care due to lack of income moderated the expected positive effect of the availability of health care services in cities for adults living in cluster 4. However, further investigations are needed to understand the unexpected negative correlation between poverty and adult mortality in this cluster. It should be kept in mind that a model is a hypothesis about how to explain the data, and the merits of alternative models often must be considered (Garfinkel et al. 2017). Despite the relevance of including contextual factors when modelling spatial patterns of mortality, the exact extent of the influences of each variable on mortality remains unknown. Moreover, no factor can be considered in isolation, because of the many interactions between death risk drivers. Several authors emphasised the complexity of the causal relationships between morbidity/mortality and socioeconomic and environmental conditions (Lucas et McMichael 2005; Goldman 2006; Henry et Dos Santos 2013). Such complexity stems from the many mediated and moderated effects through which contextual factors act on mortality. Thus, although this study provides useful indications on mortality level differentials, and their associated determinants, a greater attention must be taken when attempting to draw conclusions on mortality causes (Henry et Dos Santos 2013; Levin 1992).

Still, our study bears a few limitations. Despite the high effort to collect complete and reliable information, census data suffer from deficiencies which could bias the interpretation of results. Such deficiencies primarily relate to possible omission and misreporting of deaths and age at deaths. Furthermore, we didn't isolate infantile deaths (0-1 year) in the classification of deaths into age groups because such an approach would have led to disproportionately high infantile and/or child-adolescent crude death rates in some neighbourhoods where related population denominators are low. However, this choice may have an impact on results because the primary infant health determinants are genetic, while determinants of mortality in other age-group emphasises more the role of socioeconomic and environmental conditions. Thus, caution is needed in the interpretation of child-adolescent mortality here, as infantile mortality is included in this age group. Nevertheless, we still think that our results are valid since they provide a detailed, disaggregated and contextualised analysis of spatial patterns of mortality. Future studies could investigate trends and determinants of neonatal (< 1 month), infantile (0-1 year) and post-neonatal deaths (1-59 months), but would require more efforts to collect reliable information for such specific analyses. In addition, more sophisticated methods such as Bayesian hierarchical models would be required to deal with the issue of extreme values when working at a fine spatial scale.

Chapter 5

Spatial analysis of COVID-19 in an urban African context using geolocated individual data: the case of Dakar, Senegal[7]

[7] Adapted from : Gadiaga, A. N., Diène, A. N., Speybroeck, N., Ndiaye, N. M., Dia, I. K., Linard, C.: **Spatial analysis of COVID-19 in an urban African context using geolocated individual data: the case of Dakar, Senegal.** In preparation.

The previous chapter was concerned with a geography of mortality. A GWR modelling technique has been used to explore the varying spatial relationship between age-specific mortality rate and identified contextual risk factors, at neighbourhood level.

This chapter complete the analysis of states of health in Dakar, by focusing on the geography of a disease. Firstly, the spatial distribution of COVID-19 infection is examined through a cluster analysis. Secondly, the role of population density and variable related to population mobility on the spreads of COVID-19 infection measured using two modelling approaches: a boosted regression tree at individual level and generalized linear regression at aggregated level. Thirdly, the performance two models is compared.

Abstract

Understanding intra-urban variations of COVID-19 infections and predicting the COVID-19 spread across space is a critical step towards effective public health strategies. In low and middle-income countries, however, the scarcity of spatially disaggregated COVID-19 data often limits the COVID-19 burden analysis at broad-scale country level, and therefore raises the issue of ecological fallacy. In this study, we used point-based individual COVID-19 geolocated data in Dakar, and examined the role of population density and connectivity in their spatial distribution. We used administrative, commercial and service areas as a proxy of connectivity and distinguished population density per residential area from population density per total areas in order to capture both the living conditions (i.e. household sizes) and the neighbourhood-level effect of population density (i.e. the amount of green or empty spaces). Boosted regression tree models based on individual COVID-19 data were used to explain and predict the spatial distribution of COVID-19 incidence rates between March and June 2020. We compared the results with a binomial regression model using aggregated neighbourhood-level data. The individual-level boosted regression tree model showed a better predictive performance and highlighed an association between COVID-19 infection and both population density and connectivity. The obtained risk map highlights a higher probability of COVID-19 infection in densely populated suburbs. Measuring the role of contextual risk factors and mapping the at-risk areas could provide useful insights to policymakers for more targeted interventions.

The publication of this chapter is pending approval by the National Health Research Ethics Committee and has therefore been removed from the present edition.

Chapter 6

Conclusion

6.1 Synthesis

Health inequalities in Senegal are very pronounced and globally oppose the richest Dakar region and the rest of the country. The health indicators used so far to study these inequalities are often biased, aggregated, and not related to contextual risk factors. The relative health advantage of Dakar is expected to be unevenly distributed within the region and a spatial analysis at detailed spatial level should uncover these intra-urban variations as well as their determinants. This spatial analysis requires detailed and disaggregated health indicators. Variables related to contextual risk factors should also be resolved at disaggregated spatial level. In this thesis, we analysed spatial inequalities of health in Dakar, focusing on two health indicators: (1) crude mortality rate and age-specific (child-adolescent, adult and old age) crude mortality rates at neighbourhood level; and (2) COVID-19 infection at individual and neighbourhood levels.

Existing health data sources were first examined to assess their potential for spatial analyses. Estimation of reliable crude mortality rates by administrative units was not possible for civil registration and DHIS 2 databases due to dispcrepancies and incompletness in recording and compiling deaths. Only census data allow direct estimates of crude mortality and age-specific mortality rates, while fully covering our study area. Based on the combination of census and remote sensing data, the typology of neighbourhoods revealed the heterogeneity of the urban landscape, and identified four housing quality profiles: spontaneous settlements, low-densely populated spontaneous settlements, residential neighbourhoods with ACS areas, and whole residential neighbourhoods. Overall, variations in health status follow these variations in housing quality profiles, with higher crude mortality rates significantly associated with spontaneous settlements and lower crude mortality rates positively correlated with residential neighbourhoods. Ultimately, poor housing quality has a negative impact on health. The geography of mortality shows a heterogeneous distribution that varies according to age-specific group. Highest excess adult mortality has been particularly found in suburban and eastern peripheral areas. The spatial distribution of excess mortality for children and adolescents is quite similar to those of adults, while excess mortality for the elderly is located both in the central Dakar and the periphery. The contextual factors that explain theses spatial heterogeneities at neighbourhood level are population density, poverty, education, lack of access to health care due to a lack of income, sanitation, and access to safe water sources for drinking. Population density seems to be crucial, as it appeared significanlty and positively correlated to mortality whatever the age-group considered. Some exceptions exist, when accounting for the spatial autocorrelation of the distribution of mortality across neighbourhoods, as the magnitude of influence of contextual factors on age-specific mortality varies spatially. As with mortality, population density is also a contextual risk factor associated with the COVID-19 infection. In addition, the proportion of administrative, commercial and service areas, used here as a proxy for connectivity, is an important driver of COVID-19 infectionCOVID-19 infection rates are spatially clustered and predicted to be higher in the western part of the region.

6.2 Main findings

The general research hypothesis that underlined this thesis was:

Inequalities in states of health in Dakar are the consequence of contextual factors, including spatial variations in density, land cover and land use, as well as spatial inequalities in housing quality.

This general hypothesis has been broken down into four specific hypotheses:

- Available data are of sufficient quality to analyse health inequalities in Dakar
- Dakar is characterised by health inequalities that are to large extent related to inequalities in living conditions
- Determinants of mortality vary both across space and across age group.

6.2.1 Available data are of sufficient quality to analyse health inequalities in Dakar

In Chapter 2, we assessed the potential of available data for the analysis of health inequalities in Dakar. There is a significant gap between data recorded at the civil registration offices and the DHIS 2. In 2016, for example, deaths registered at the civil registration offices were twice as high as those notified in the DHIS 2 platform. Overall, information is incomplete and often fragmented in the two data sources mainly due to an irregular data transmission process from secondary civil registration centres to MCRC on the one hand, and from health districts to the DHIS 2 platform, on the other hand. This prevents the obtention of reliable estimates of mortality rates at the relevant spatial level for this research. Although the higher volume of records in the civil registration data, the extent of the MCRC database is far from covering the whole study area. Such deficiencies were overcome by the census data. Census data is the only spatially representative database that can be used to provide mortality data of the whole study area and that can be helpful in conducting spatial analyses. The census data have been used to analyse relationships between housing quality and crude mortality rate (Chapter 3) and in modelling of the role of contextual factors in the spatial variations of age-specific mortality rates (Chapter 4). As a consequence, the first research hypothesis: **Available data are of sufficient quality to analyse health inequalities in Dakar**, is verified.

6.2.2 Dakar is characterised by health inequalities that are to large extent related to inequalities in living conditions

A hierarchical classification based on a principal component analysis revealed the heterogeneity of the Dakar's neighbourhoods in terms of living conditions. Land cover and land uses types, housing characteristics and population density were derived from census and very-high resolution satellite imagery in order to characterize living conditions in Dakar. The analysis revealed two profiles of peri-urban neighbourhoods: the dense spontaneous settlements in the suburb of Pikine, and the spontaneous neighbourhoods on the eastern front line of urbanization, characterized by low density of population . Residential planned neighbourhoods are characterized by a higher level of comfort, which translates into a better health, as a lower crude mortality rates were observed in residential neighbourhoods compared to spontaneous settlements. Chapter 4 and 5 provided a deeper understanding of health inequalities related to inequalities in living conditions- Crude mortality rates, age-specific mortality rates (Chapter 4) and cumulated incidence of COVID-19 (Chapter 5) were the health indicators analysed. Local spatial autocorrelation tests revealed existence of clusters for both these health indicators. GWPR and BRT models were used in measuring the role of contextual factors in age-specific spatial variation mortality and COVID-19 infection, respectively . Predicted spatial distribution of COVID-19 incidence showed areas at high- risk in the western parts of Dakar, particularly the areas marked by high population density and high presence of non-redisential ACS built-up areas. As a consequence, the second specific hypothesis: **Dakar is characterised by health inequalities that are to a large extent related to inequalities in living conditions**, is verified.

6.2.3 Determinants of mortality vary both across space and across age group

Previous studies analysing determinants of mortality suspected a higher impact of environmental determinants in the lowest age-groups, because child mortality is dominated by infectious and parasitic diseases, whereas a higher impact of health behaviours for adult and elderly. Here socioeconomic and environmental contextual risk factors were related to age-specific mortality rates in order to identify age-specific health determinants in Dakar. Results show a consistent impact of population density, whatever the age group or the location in the city. The set of variables involved in explaining spatial variations of child-adolescent mortality is slightly different to those involved in determining adult mortality, and includes population density, education, poverty, sanitation, lack of access to health care services due to lack of income and access to clean water for drinking. Regarding old age mortality, only population density was a significant determinant, except for access to safe water source for drinking for a few neighbourhoods. Therefore, the third hypothesis: **Determinants of mortality vary both across space and across age group,** is verified.

The general hypothesis of this thesis research is confirmed through the validation of the three specific research hypotheses. Our research results are robust to the various indicators used to reflect the states of health and living conditions.

6.3 Scientific contribution

Returning to the conceptual model (Figure 1.1), states of health are conceived as the result of the interplay between population, habitat and behaviour. Also, causal relationships are complex due to the issue of scale and the risk of ecological fallacy. In line with these various concepts, the three following scientific contributions of the thesis can be highlighted.

6.3.1 In-depth analysis of the determinants of mortality

A major scientific contribution of this thesis is the decomposition of the crude mortality rate into age-specific mortality rates, which allow an in-depth analysis of the determinants of mortality. Altough the long history of measuring determinants of mortality in Africa, the majority of these studies chiefly focused on child mortality and there is a striking scarcity of studies examining determinants of adult mortality. To our knowledge, this is the first study that analyse spatial inequalities in health through the examination of child-adolescent, adult and old age mortality.

The prevalence, incidence and determinants of mortality are age- and place- specific. Measuring how the effects of the contextual variables on age-specific mortality vary spatially provided a perfect illustration of the in-depth analysis of the determinants of mortality conducted in this research. Indeed, the uneven distribution of deaths across space (shown by existence of spatial autocorrelation) follows an unequal spatial distribution of their determinants. Depending on the neighbourhood or place where one live in the city, negative and positive influence are cumulating to reflect the health status (WHO - UN Habitat 2009). Place of residence influences the risk of death, as already shown by Gruebner et al. (2015) in Kenya, where living in non-slum urban areas was associated with significantly lower infant death rates as compared to living in rural and informal settlements. In this research, the relative contribution of health risk factors was assessed through cluster-specific regression analyses, which capture spatial differences in the influences of the determinants. While living in urban areas is generally associated with a better health than in rural areas, such an advantage may be moderated by a lack of access to health care due to a lack of income or a low education level.

6.3.2 Assessment of the potential of health statistics

The completeness of death registrations is high in the Dakar region compared to other regions in Senegal. According to the last census (2013), 88% of deaths were registered in the civil registration offices (while it was 80% in 2002), compared to only 30% at the national level (RGPHAE-2013). Here, the high completeness of civil registration data was also confirmed by comparing DHIS 2 data and deaths declared at the civil registration offices. Civil registration could allow a better monitoring of demographic and health changes. However, several issues need to be tackled first. There is a high demand in the pre-processing of the data. Analysis of spatial distribution of apparent mortality rates and seasonal patterns of deaths, revealed potential of civil registration in monitoring continuities and changes of health trends. Cluster analysis was therefore restricted to the department of Dakar, due to several discrepancies, especially in the variables related to - place of deaths.

6.3.3 Combination of census and surveys with remote sensing data

The increased availability of remotely sensed data raised issues about the usefulness in combining remote sensing with census and survey data for health studies. Combining census and remote sensing data provided a better characterisation of the urban landscape and analysis of spatial inequalities of living conditions in Dakar (Borderon et al. 2014). It can be concluded from this analysis that census and remote sensing data are very helpful in understanding health determinants in different geographic areas, as inequalities in living conditions are proxies for differences in health status. In the modelling of infectious diseases infection and spread, remote sensing data play a crucial role in identifying pathogens, hosts and vectors habitats preferences. Other environmental risk factors related to non-communicable diseases such as harmful substances (carbon monoxide, nitrogen oxides, sulphur dioxide and ozone) can be measured using remote sensing technologies (Jia et Stein 2017). In this thesis, census and land cover and land use products were combined to characterize housing quality and its impact on health at the neighbourhood level. In addition, the spatial modelling of the COVID-19 infection in Dakar through a combination of remote sensing and survey data is another noticeable scientific contribution, as land cover and land use data were shown to be important risk factors.

6.4 Methodological contributions

6.4.1 The need for fine-scale georeferenced census data

Given the imperfection of health data sources such as the civil registration system or DHIS 2, censuses remain the preferred source for understanding states of health and their determinants, despite their low frequency (i.e. every 10 years). However, to be used to analyse intra-urban inequalities, census data must be available at a fine spatial scale. Significant efforts have been made in the framework of this doctoral research to integrate the 2013 census database into a GIS system by precisely delineating neigbourhood boundaries for the Dakar region.

6.4.2 Dealing with spatial autocorrelation and overdispersion

Another important methodological contribution of the thesis relates to the spatial modelling approaches used. The applicability of spatial regression models largely depends on data availability and quality. In many instances, surveillance systems and other health data sources do not provide data of sufficient quality to build the best-suited model. Missing values, zero-inflated distributions and data scarcity are major bottlenecks for applying robust spatial statistical inferences in health studies. In addition, geolocated epidemiological data often suffer from spatial dependency problems. Different strategies have been successfully applied in this thesis in order to deal with both spatial autocorrelation and overdispersion problems detected: (i) in Chapter 4, GWPR and clusters-specific quasi Poisson regression models were used to simultaneously handle the spatial autocorrelation and the overdispersion detected in the age-specific mortality rates, (ii) BRT have been used in Chapter 5 in order to predict the spatial distribution of COVID-19 incidence rates based on individual data. Such machine learning techniques have proven to be particularly useful to handle spatial autocorrelation and overdispersion issues.

6.4.3 Choosing the right scale of analysis

A third methodological contribution of this thesis is the comparison of the predictive performance of models based on individual and aggregated data (Chapter 5). Our results showed a better predictive performance when using individually-based COVID-19 infections than aggregated incidence rates. In addition, there is a higher risk of ecological fallacy when using aggregated incidence rates, even if the risk is also present in our individual-based model given that risk factors were summarized for buffer zones around each infection point. Choosing the most appropriate scale of analysis can be a challenge, especially as the finest is not always the best and fine-scale health data may raise privacy concerns. Ultimately, the optimal scale of analysis rely on the health problem under consideration.

6.5 Methodological limitations

6.5.1 Data availability and quality

This thesis began with a quality assessment of existing health data as well as their representativeness for spatial analyses. The potential contribution of civil registration data to understand spatial health inequalities in Dakar is hampered by the scattered nature of the data. To be useful for such analyses, the civil registration system must be improved by digitizing records and making sure that data are transferred from secondary centres to the main civil registration centre. Efforts in that direction are underway but currently insufficient.

Although efforts were made to collect and assemble reliable information, the health data analysed here suffer from several deficiencies that could lead to biased interpretations of results. Census data may contain errors such as death omissions or misreporting of dates and ages at death. Possible misreporting of age at death may alter the spatial distribution of age-specific mortality rates analysed in Chapter 4. In Chapter 5, the extent of the COVID-19 burden is not fully assessed because data used only concern infected and then recovered patients. Nothing is known about COVID-19 deceased patients. In addition, the sample only includes COVID-19 confirmed cases (i.e. patients that have been tested), which most probably largely under-estimate the real number of cases and call for caution when interpreting the results, especially the assessment of the disease extent.

6.5.2 Spatio-temporal analysis

It should also be noted that the temporal dimension of health indicators has often been overlooked. The census-derived mortality data (Chapters 3 and 4) and the COVID-19 survey data (Chapter 5) do not allow the analysis of trends over time. Seasonal patterns are however potential drivers of many diseases, with a direct and/or indirect impact on disease burden and subsequent deaths (Henry et Dos Santos 2013), such as transmission dynamics depending on environmental and social factors for example. The local clustering effects modelled in Chapters 4 and 5 are based on discrete and cumulated counts over a certain period of time, which means that the spatial autocorrelation was not tested in the temporal domain.

Beyond the lack of temporal information in health data, contextual information used in the thesis such as satellite-derived land cover and land use data are only available for a single date. Much more time and ressources are needed to produce time-series that would allow to take urban changes into account.

6.6 Perspectives for further research

6.6.1 Methodological research

A major improvement that could be made to the spatial approach developed in this thesis is enhancing models with an analysis capability in the temporal domain. The census-derived neighbourhood housing quality indices developed in Chapter 3 could be applied to the 2002, or the upcoming 2023 censuses in order to study the evolution of health profiles through time. The same is true for the analysis of the age-specific mortality determinants (Chapter 4). In addition, these age-specific mortality determinants could be analysed using standardised mortality ratios calculated using Bayesian spatial models in order to avoid the issue of extreme values in sparsely-populated neighbourhoods (i.e. by calculating expected counts for each neighbourhood). Besides, Bayesian spatial models may be particularly useful in modelling the determinants of the geographic distribution of mortality as they include a spatial random effect that smoothes the data according to the neighbourhood structure (Moraga, 2019). Lastly, spatial hierarchical models merit attention in future studies in order to improve the reliability of the scale used for analysis.

6.6.2 Applications

This research raised the potential of census data for analysing spatial health inequalities at the intra-urban level and could help to bridge the gap currently existing regarding such fine-scale analyses in low-and middle-income countries. The spatial approach developed here could be applied to other African cities, particularly those that have recently conducted a census using location-based devices for data collection and GIS tools for the creation of census tracts. There are clear improvements opportunities in terms of data collection, visualization, processing and interpretation of health-related census indicators. In addition, the BRT approach based on individual data presented in Chapter 5 for predicting the risk of COVID-19 infection could support the development of a disease surveillance system for the ongoing COVID-19 pandemic, for other major endemic diseases (such as malaria and tuberculosis) or for future epidemies. Such approaches can however be sensitive as they are based on individual health data and should therefore be framed by clear privacy protection rules.

Finally, this thesis is an excellent illustration of the usefulness of interdisciplinarity in sciences. Achieving our research objectives has been possible through close collaboration between geographers, demographers, public health practitioners and biologists, each having brought their knowledge and experience in spatial and statistical analysis and modelling, health determinants, and predictive analyses. Another useful application of this research is the replicability of the methods used here by these stakeholders in future studies.

6.7 Implications for policy

Identifying at-risk areas in addition to at-risk population groups, as has been done here, has important implications for public health policies. This doctoral research has provided a detailed description of health states in the Dakar region, with a particular focus on living conditions and other social and environmental health determinants. Geo-visualisation and interpretation of the results obtained can be useful resources in assisting the implementation of upstream and downstream public health interventions. Specifically, the spatial analysis of age-specific mortality has given quantitative and qualitative information about the relative exposure of neighbourhoods to mortality risk. Such information provides valuable guidelines for geographically targeted interventions, e.g. the optimized distribution of health care facilities, the treatment of infectious diseases affecting more the younger generations or chronic diseases affecting mostly older generations.

At the beginning of the COVID-19 pandemic in 2020, the Senegalese ministry of health deployed an online platform for mapping and monitoring the daily confirmed COVID-19 cases at health district and municipality levels, without using any spatial analysis tool to take full advantage of the data collected. In the following months, spatial tools such as cluster detection and BRT models have however proven to be very useful as decision support tools. There is an urgent need to further develop such GIS and modelling techniques and to transfer knowledge to low- and middle-income countries is an urgent need, especially at a time when vaccination campaigns need to be organized and priority areas targeted.

The degree of challenge in exploring spatial dynamics is exaggerated by the scarcity of data (Wen et al. 2014). Improving health surveillance systems is needed to meet the data requirements of spatial models analysing health determinants. Reducing the risk of overdispersion through data quality enhancement would be quite more convenient than using alternative regression models.

Bibliography

Adjei, Prince Osei-Wusu, et Peter Ohene Kyei. 2013. « Linkages between Income, Housing Quality and Disease Occurrence in Rural Ghana ». *Journal of Housing and the Built Environment* 28 (1): 35-49. https://doi.org/10.1007/s10901-012-9277-6.

Aikins, Ama de-Graft, Nigel Unwin, Charles Agyemang, Pascale Allotey, Catherine Campbell, et Daniel Arhinful. 2010. « Tackling Africa's chronic disease burden: from the local to the global ». *Globalization and Health* 6 (1): 5. https://doi.org/10.1186/1744-8603-6-5.

Alirol, Emilie, Laurent Getaz, Beat Stoll, François Chappuis, et Louis Loutan. 2011. « Urbanisation and infectious diseases in a globalised world ». *The Lancet Infectious Diseases* 11 (2): 131-41. https://doi.org/10.1016/S1473-3099(10)70223-1.

Anselin, Luc. 1995. « Local Indicators of Spatial Association—LISA ». *Geographical Analysis* 27 (2): 93-115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

Arias, Elizabeth, et Susan De Vos. 1996. « Using Housing Items to Indicate Socioeconomic Status: Latin America ». *Social Indicators Research* 38 (1): 53-80. https://doi.org/10.1007/BF00293786.

Arthur, Ronan F., Emily S. Gurley, Henrik Salje, Laura S. P. Bloomfield, et James H. Jones. 2017. « Contact Structure, Mobility, Environmental Impact and Behaviour: The Importance of Social Forces to Infectious Disease Dynamics and Disease Ecology ». *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 372 (1719). https://doi.org/10.1098/rstb.2016.0454.

Ayubi, Erfan, et Saeid Safiri. 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 ». *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association* 27 (4): 1123. https://doi.org/10.1016/j.jstrokecerebrovasdis.2017.11.039.

Balk, Deborah, Thomas Pullum, Adam Storeygard, Fern Greenwell, et Melissa Neuman. 2004. « A Spatial Analysis of Childhood Mortality in West Africa ». *Population, Space and Place* 10 (3): 175-216. https://doi.org/10.1002/psp.328.

Barnett, Adrian G., et Annette J. Dobson. 2010. « Introduction ». In *Analysing Seasonal Health Data*, édité par Adrian G. Barnett et Annette J. Dobson, 1-47. Statistics for Biology and Health. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-10748-1_1.

Bawah, Ayaga A. & Zuberi. s. d. « Socioeconomic Status and Child Mortality: An Illustration Using Housing and Household Characteristics from African Census Data ». *African Population Studies* 19 (SB): 9-29.

Beck, L. R., B. M. Lobitz, et B. L. Wood. 2000. « Remote sensing and human health: new sensors and new opportunities. » *Emerging Infectious Diseases* 6 (3): 217-27.

Berke, Olaf. 2004. « Exploratory disease mapping: kriging the spatial risk function from regional count data ». *International Journal of Health Geographics* 3 (1): 18. https://doi.org/10.1186/1476-072X-3-18.

Bernard, Paul-Marie, et Claude Lapointe. 1987. *Mesures statistiques en épidémiologie*. PUQ.

Bhadra, Arunava, Arindam Mukherjee, et Kabita Sarkar. 2021. « Impact of Population Density on Covid-19 Infected and Mortality Rate in India ». *Modeling Earth*

Systems and Environment 7 (1): 623-29. https://doi.org/10.1007/s40808-020-00984-7.
Bhatta, B., S. Saraswati, et D. Bandyopadhyay. 2010. « Urban sprawl measurement from remote sensing data ». *Applied Geography*, Climate Change and Applied Geography – Place, Policy, and Practice, 30 (4): 731-40. https://doi.org/10.1016/j.apgeog.2010.02.002.
Boadi, Kwasi, Markku Kuitunen, Kolawole Raheem, et Kari Hanninen. 2005. « Urbanisation Without Development: Environmental and Health Implications in African Cities ». *Environment, Development and Sustainability* 7 (4): 465-500. https://doi.org/10.1007/s10668-004-5410-3.
Borderon, Marion, et Sébastien Oliveau. 2017. « Vulnérabilités sociales et changement d'échelle. L'espace compte ! » *Espace populations sociétés. Space populations societies*, nº 2016/3 (janvier). https://doi.org/10.4000/eps.7012.
Borderon, Marion, Sébastien Oliveau, Vanessa Machault, Cécile Vignolles, Jean-Pierre Lacaux, et Alphousseyni N'Donky. 2014. « Qualifier les espaces urbains à Dakar, Sénégal. Résultats préliminaires de l'approche croisée entre télédétection et données censitaires spatialisées ». *Cybergeo : European Journal of Geography*, mars. https://doi.org/10.4000/cybergeo.26250.
Boyce, Matthew R., Rebecca Katz, et Claire J. Standley. 2019. « Risk Factors for Infectious Diseases in Urban Environments of Sub-Saharan Africa: A Systematic Review and Critical Appraisal of Evidence ». *Tropical Medicine and Infectious Disease* 4 (4). https://doi.org/10.3390/tropicalmed4040123.
Carozzi, Felipe. 2020. « Urban Density and Covid-19 ». SSRN Scholarly Paper ID 3643204. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3643204.
Caselli, Graziella. s. d. « Les entorses au schéma de la transition épidémiologique ». Consulté le 6 janvier 2022. https://www.academia.edu/26466078/Les_entorses_au_sch%C3%A9ma_de_la_transition_%C3%A9pid%C3%A9miologique.
Chamberlain, Heather, Attila Lazar, et Andrew Tatem. 2021. « Ease of Social Distancing Index, Mapped for Urban Areas in Sub-Saharan Africa, Version 1.0 ». University of Southampton. https://doi.org/10.5258/SOTON/WP00711.
Chen, Dongmei. 2014. « Modeling the Spread of Infectious Diseases: A Review ». In *Analyzing and Modeling Spatial and Temporal Dynamics of Infectious Diseases*, 19-42. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118630013.ch2.
Chen, Dongmei, Bernard Moulin, et Jianhong Wu. 2014. *Analyzing and Modeling Spatial and Temporal Dynamics of Infectious Diseases*. John Wiley & Sons.
Chinazzi, Matteo, Jessica T. Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, et al. 2020. « The Effect of Travel Restrictions on the Spread of the 2019 Novel Coronavirus (COVID-19) Outbreak ». *Science* 368 (6489): 395-400. https://doi.org/10.1126/science.aba9757.
Clark, Samuel J. 2019. « A General Age-Specific Mortality Model With an Example Indexed by Child Mortality or Both Child and Adult Mortality ». *Demography* 56 (3): 1131-59. https://doi.org/10.1007/s13524-019-00785-3.
« Communicable Diseases ». s. d. WHO | Regional Office for Africa. Consulté le 13 octobre 2021. https://www.afro.who.int/health-topics/communicable-diseases.
Comrey, Andrew L., et Howard B. Lee. 2016. *A First Course in Factor Analysis*. 2 edition. Routledge.

Cordes, Jack, et Marcia C. Castro. 2020. « Spatial Analysis of COVID-19 Clusters and Contextual Factors in New York City ». *Spatial and Spatio-Temporal Epidemiology* 34 (août): 100355. https://doi.org/10.1016/j.sste.2020.100355.

« Creating spatial interpolation surfaces with DHS data - ePrints Soton ». s. d. Consulté le 19 décembre 2019. https://eprints.soton.ac.uk/382342/.

Cressie, Noel. 2015. *Statistics for Spatial Data*. John Wiley & Sons.

Cutler, David, Angus Deaton, et Adriana Lleras-Muney. 2006. « The Determinants of Mortality ». *Journal of Economic Perspectives* 20 (3): 97-120. https://doi.org/10.1257/jep.20.3.97.

« Deriving fine-scale socioeconomic information of urban areas using very high-resolution satellite imagery: International Journal of Remote Sensing: Vol 32, No 21 ». s. d. Consulté le 19 décembre 2019. https://www.tandfonline.com/doi/abs/10.1080/01431161.2010.512928.

Diallo, Abdoulaye, Nicaise Tuikue Ndam, Azizath Moussiliou, Stéphanie Dos Santos, Alphousseyni Ndonky, Marion Borderon, Sébastien Oliveau, Richard Lalou, et Jean-Yves Le Hesran. 2012. « Asymptomatic Carriage of Plasmodium in Urban Dakar: The Risk of Malaria Should Not Be Underestimated ». *PLOS ONE* 7 (2): e31100. https://doi.org/10.1371/journal.pone.0031100.

Diez-Roux, A V. 1998. « Bringing context back into epidemiology: variables and fallacies in multilevel analysis. » *American Journal of Public Health* 88 (2): 216-22. https://doi.org/10.2105/AJPH.88.2.216.

Dlamini, Sabelo Nick, Anton Beloconi, Sizwe Mabaso, Penelope Vounatsou, Benido Impouma, et Ibrahima Socé Fall. 2019. « Review of Remotely Sensed Data Products for Disease Mapping and Epidemiology ». *Remote Sensing Applications: Society and Environment* 14 (avril): 108-18. https://doi.org/10.1016/j.rsase.2019.02.005.

Dlodlo, Riitta A., Paula I. Fujiwara, Zanele E. Hwalima, Stanley Mungofa, et Anthony D. Harries. 2011. « Adult Mortality in the Cities of Bulawayo and Harare, Zimbabwe: 1979-2008 ». *Journal of the International AIDS Society* 14 (S1): S2-S2. https://doi.org/10.1186/1758-2652-14-S1-S2.

Doctor, Henry V. 2004. « The Effect of Living Standards on Childhood Mortality in Malawi », décembre. https://tspace.library.utoronto.ca/handle/1807/5820.

Dos Santos, Stéphanie, Iulia Rautu, Mody Diop, Mahaman Mourtala Abdou Illou, Alphousseyni Ndonky, Jean-Yves Le Hesran, et Richard Lalou. 2015. « The Influence of Environmental Factors on Childhood Fever during the Rainy Season in an African City: A Multilevel Approach in Dakar, Senegal ». *Population and Environment* 36 (4): 429-51. https://doi.org/10.1007/s11111-014-0224-1.

Duboz, P., L. Gueye, G. Boetsch, et E. Macia. 2015. « Access to health care in Dakar (Senegal): frequency, type of provider, and non-communicable chronic diseases ». *Médecine et Santé Tropicales* 25 (2): 165-71. https://doi.org/10.1684/mst.2015.0450.

Dye, Christopher. 2008. « Health and Urban Living ». *Science (New York, N.Y.)* 319 (5864): 766-69. https://doi.org/10.1126/science.1150198.

El Bcheraoui, Charbel, Honoré Mimche, Yodé Miangotar, Varsha Sarah Krish, Faye Ziegeweid, Kris J Krohn, Martin Herbas Ekat, et al. 2020. « Burden of Disease in Francophone Africa, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017 ». *The Lancet Global Health* 8 (3): e341-51. https://doi.org/10.1016/S2214-109X(20)30024-3.

Emch, Michael, Elisabeth Dowling Root, et Margaret Carrel. 2017. *Health and Medical Geography, Fourth Edition*. Guilford Publications.
Fann, Rocade. 2014. « AGENCE NATIONALE DE LA STATISTIQUE ET DELA DEMOGRAPHIE », 7.
Feng, Luzhao, Ting Zhang, Qing Wang, Yiran Xie, Zhibin Peng, Jiandong Zheng, Ying Qin, et al. 2021. « Impact of COVID-19 Outbreaks and Interventions on Influenza in China and the United States ». *Nature Communications* 12 (1): 3249. https://doi.org/10.1038/s41467-021-23440-1.
Ferreira, Marcos César. 2020. « Spatial Association between the Incidence Rate of COVID-19 and Poverty in the São Paulo Municipality, Brazil ». *Geospatial Health* 15 (2). https://doi.org/10.4081/gh.2020.921.
Fotso, Jean-Christophe, et Barthelemy Kuate-Defo. 2005. « Measuring Socioeconomic Status in Health Research in Developing Countries: Should We Be Focusing on Households, Communities or Both? » *Social Indicators Research* 72 (2): 189-237. https://doi.org/10.1007/s11205-004-5579-8.
Gadiaga, Assane, Florence de Longueville, Stefanos Georganos, Taïs Grippa, Sébastien Dujardin, Diene Aminata, Bruno Masquelier, Mouhamadou Diallo, et Catherine Linard. 2021. « Neighborhood-Level Housing Quality Indices for Health Assessment in Dakar, Senegal ». *Geospatial Health*, janvier. http://hdl.handle.net/2013/.
Garfinkel, Alan, Jane Shevtsov, et Yina Guo. 2017. « Modeling, Change, and Simulation ». In *Modeling Life: The Mathematics of Biological Systems*, édité par Alan Garfinkel, Jane Shevtsov, et Yina Guo, 1-68. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-59731-7_1.
Garland, Patrick, Dave Babbitt, Maksym Bondarenko, Alessandro Sorichetta, Andrew J. Tatem, et Oliver Johnson. 2020. « The COVID-19 pandemic as experienced by the individual ». *arXiv:2005.01167 [physics, q-bio, stat]*, septembre. http://arxiv.org/abs/2005.01167.
Georganos, Stefanos, Abdulhakim M. Abdi, David E. Tenenbaum, et Stamatis Kalogirou. 2017. « Examining the NDVI-Rainfall Relationship in the Semi-Arid Sahel Using Geographically Weighted Regression ». *Journal of Arid Environments* 146 (novembre): 64-74. https://doi.org/10.1016/j.jaridenv.2017.06.004.
Georganos, Stefanos, Oscar Brousse, Sébastien Dujardin, Catherine Linard, Daniel Casey, Marco Milliones, Benoit Parmentier, et al. 2020. « Modelling and mapping the intra-urban spatial distribution of Plasmodium falciparum parasite rate using very-high-resolution satellite derived indicators ». *International Journal of Health Geographics* 19 (1). https://doi.org/10.1186/s12942-020-00232-2.
Georganos, Stefanos, Assane Niang Gadiaga, Catherine Linard, Tais Grippa, Sabine Vanhuysse, Nicholus Mboga, Eléonore Wolff, Sébastien Dujardin, et Moritz Lennert. 2019a. « Modelling the Wealth Index of Demographic and Health Surveys within Cities Using Very High-Resolution Remotely Sensed Information ». *Remote Sensing* 11 (21): 2543. https://doi.org/10.3390/rs11212543.
———. 2019b. « Modelling the Wealth Index of Demographic and Health Surveys within Cities Using Very High-Resolution Remotely Sensed Information ». *Remote Sensing* 11 (21): 2543. https://doi.org/10.3390/rs11212543.
Gilbert, Marius, Giulia Pullano, Francesco Pinotti, Eugenio Valdano, Chiara Poletto, Pierre-Yves Boëlle, Eric D'Ortenzio, et al. 2020. « Preparedness and Vulnerability of African Countries against Importations of COVID-19: A Modelling Study ».

Lancet (London, England) 395 (10227): 871-77. https://doi.org/10.1016/S0140-6736(20)30411-6.
Goldman, Noreen. 2006. « Social Inequalities in Health: Disentangling the Underlying Mechanisms ». *Annals of the New York Academy of Sciences* 954 (1): 118-39. https://doi.org/10.1111/j.1749-6632.2001.tb02750.x.
Gomes, D. S., L. A. Andrade, C. J. N. Ribeiro, M. V. S. Peixoto, S. V. M. A. Lima, A. M. Duque, T. M. Cirilo, et al. 2020. « Risk Clusters of COVID-19 Transmission in Northeastern Brazil: Prospective Space–Time Modelling ». *Epidemiology & Infection* 148. https://doi.org/10.1017/S0950268820001843.
Grippa, Tais, et Stefanos Georganos. 2018. « Dakar Land Use Map At Street Block Level ». Zenodo. https://doi.org/10.5281/zenodo.1291389.
Grippa, Tais, Stefanos Georganos, Sabine G. Vanhuysse, Moritz Lennert, et Eléonore Wolff. 2017. « A local segmentation parameter optimization approach for mapping heterogeneous urban environments using VHR imagery ». In *Remote Sensing Technologies and Applications in Urban Environments II*, édité par Wieke Heldens, Nektarios Chrysoulakis, Thilo Erbertseder, et Ying Zhang, 20. Warsaw, Poland: SPIE. https://doi.org/10.1117/12.2278422.
Grippa, Taïs, Stefanos Georganos, Soukaina Zarougui, Pauline Bognounou, Eric Diboulo, Yann Forget, Moritz Lennert, Sabine Vanhuysse, Nicholus Mboga, et Eléonore Wolff. 2018a. « Mapping Urban Land Use at Street Block Level Using OpenStreetMap, Remote Sensing Data, and Spatial Metrics ». *ISPRS International Journal of Geo-Information* 7 (7): 246. https://doi.org/10.3390/ijgi7070246.
———. 2018b. « Mapping Urban Land Use at Street Block Level Using OpenStreetMap, Remote Sensing Data, and Spatial Metrics ». *ISPRS International Journal of Geo-Information* 7 (7): 246. https://doi.org/10.3390/ijgi7070246.
Grippa, Taïs, Moritz Lennert, Benjamin Beaumont, Sabine Vanhuysse, Nathalie Stephenne, et Eléonore Wolff. 2017. « An Open-Source Semi-Automated Processing Chain for Urban Object-Based Classification ». *Remote Sensing* 9 (4): 358. https://doi.org/10.3390/rs9040358.
Grippa, Taïs, Catherine Linard, Moritz Lennert, Stefanos Georganos, Nicholus Mboga, Sabine Vanhuysse, Assane Gadiaga, et Eléonore Wolff. 2019. « Improving Urban Population Distribution Models with Very-High Resolution Satellite Information ». *Data* 4 (1): 13. https://doi.org/10.3390/data4010013.
Gruebner, Oliver, Sven Lautenbach, M. M. H. Khan, Samuel Kipruto, Michael Epprecht, et Sandro Galea. 2015. « Place of Residence Moderates the Risk of Infant Death in Kenya: Evidence from the Most Recent Census 2009 ». *PLOS ONE* 10 (10): e0139545. https://doi.org/10.1371/journal.pone.0139545.
Gulyani, Sumila, Ellen M. Bassett, et Debabrata Talukdar. 2014. « A Tale of Two Cities: A Multi-Dimensional Portrait of Poverty and Living Conditions in the Slums of Dakar and Nairobi ». *Habitat International* 43 (juillet): 98-107. https://doi.org/10.1016/j.habitatint.2014.01.001.
Günther, Isabel, et Kenneth Harttgen. 2012. « Deadly Cities? Spatial Inequalities in Mortality in Sub-Saharan Africa ». *Population and Development Review* 38 (3): 469-86. https://doi.org/10.1111/j.1728-4457.2012.00512.x.
Hamidi, Shima, Sadegh Sabouri, et Reid Ewing. 2020. « Does Density Aggravate the COVID-19 Pandemic? » *Journal of the American Planning Association* 86 (4): 495-509. https://doi.org/10.1080/01944363.2020.1777891.
Harttgen, Kenneth, Stefan Lang, Judith Santer, et Johannes Seiler. 2020. « Modelling Under-Five Mortality through Multilevel Structured Additive Regression with

Varying Coefficients for Asia and Sub-Saharan Africa ». *The Journal of Development Studies* 56 (2): 401-30. https://doi.org/10.1080/00220388.2018.1563681.
Hazarie, Surendra, David Soriano-Paños, Alex Arenas, Jesús Gómez-Gardeñes, et Gourab Ghoshal. 2021. « Interplay between intra-urban population density and mobility in determining the spread of epidemics ». *arXiv:2102.00671 [physics]*, février. http://arxiv.org/abs/2102.00671.
Henry, Sabine J. F., et Stéphanie Dos Santos. 2013. « Rainfall Variations and Child Mortality in the Sahel: Results from a Comparative Event History Analysis in Burkina Faso and Mali ». *Population and Environment* 34 (4): 431-59. https://doi.org/10.1007/s11111-012-0174-4.
Herrin, William E., Michelle M. Amaral, et Arsene M. Balihuta. 2013. « The Relationships between Housing Quality and Occupant Health in Uganda ». *Social Science & Medicine* 81 (mars): 115-22. https://doi.org/10.1016/j.socscimed.2012.12.017.
« Hidden Cities ABRIDGED English ». s. d. Consulté le 7 mai 2019. https://unhabitat.org/books/hidden-cities-abridged-english/.
« Home ». s. d. Africa CDC. Consulté le 13 octobre 2021. https://africacdc.org/.
Hussein, Asim Abdelmoneim. 2014. « The "Epidemiological Transmission" and "Double-Burden of Disease": A Focus on Africa ». *Burden of Disease* 2 (2): 9.
Iyun, B. Folasade. 1993. « The geographical inequalities in mortality in Africa ». *Social Science & Medicine*, Special Issue Geographical Equalities of Mortality in Developing Countries, 36 (10): 1243-45. https://doi.org/10.1016/0277-9536(93)90213-N.
J. Madise, Nyovani, Abdhalah K. Ziraba, Joseph Inungu, Samoel A. Khamadi, Alex Ezeh, Eliya M. Zulu, John Kebaso, Vincent Okoth, et Matilu Mwau. 2012. « Are Slum Dwellers at Heightened Risk of HIV Infection than Other Urban Residents? Evidence from Population-Based HIV Prevalence Surveys in Kenya ». *Health & Place* 18 (5): 1144-52. https://doi.org/10.1016/j.healthplace.2012.04.003.
Jacqueline, Bouyer Jean, Hémon Denis, Cordier Sylvaine, Derriennic Francis, Stücker Isabelle, Stengel Bénédicte, CLAVEL. 2009. *Epidémiologie Principes et méthodes quantitatives*. Lavoisier.
Jankowska, Marta M., Magdalena Benza, et John R. Weeks. 2013. « Estimating spatial inequalities of urban child mortality ». *Demographic research* 28 (janvier): 33-62. https://doi.org/10.4054/DemRes.2013.28.2.
Jia, Peng, et Alfred Stein. 2017. « Using remote sensing technology to measure environmental determinants of non-communicable diseases ». *International Journal of Epidemiology* 46 (4): 1343-44. https://doi.org/10.1093/ije/dyw365.
Juma, Kenneth, Pamela A. Juma, Constance Shumba, Peter Otieno, et Gershim Asiki. 2019. *Non-Communicable Diseases and Urbanization in African Cities: A Narrative Review*. *Public Health in Developing Countries - Challenges and Opportunities*. IntechOpen. https://doi.org/10.5772/intechopen.89507.
Kabaria, Caroline W., Fabrizio Molteni, Renata Mandike, Frank Chacky, Abdisalan M. Noor, Robert W. Snow, et Catherine Linard. 2016. « Mapping intra-urban malaria risk using high resolution satellite imagery: a case study of Dar es Salaam ». *International Journal of Health Geographics* 15 (1): 26. https://doi.org/10.1186/s12942-016-0051-y.
Kadi, Nadjat, et Mounia Khelfaoui. 2020. « Population Density, a Factor in the Spread of COVID-19 in Algeria: Statistic Study ». *Bulletin of the National Research Centre* 44 (1): 138. https://doi.org/10.1186/s42269-020-00393-x.

Kessides, Christine. 2007. « The Urban Transition in Sub-Saharan Africa: Challenges and Opportunities ». *Environment and Planning C: Government and Policy* 25 (4): 466-85. https://doi.org/10.1068/c3p.
Kofie, Richard Y., Emmanuel M. Attua, et John S. Nabila. 2008. « Poverty and socio-economic consequences of Buruli ulcer (Mycobacterium ulcerans) in the Ga West District of Ghana ». *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography* 62 (3): 210-21. https://doi.org/10.1080/00291950802335855.
Korah, Prosper Issahaku, Tony Matthews, et Deanna Tomerini. 2019. « Characterising Spatial and Temporal Patterns of Urban Evolution in Sub-Saharan Africa: The Case of Accra, Ghana ». *Land Use Policy* 87 (septembre): 104049. https://doi.org/10.1016/j.landusepol.2019.104049.
Lai, Shengjie, Nick W. Ruktanonchai, Alessandra Carioli, Corrine W. Ruktanonchai, Jessica R. Floyd, Olivia Prosper, Chi Zhang, Xiangjun Du, Weizhong Yang, et Andrew J. Tatem. 2021. « Assessing the Effect of Global Travel and Contact Restrictions on Mitigating the COVID-19 Pandemic ». *Engineering*, mai, S2095809921001983. https://doi.org/10.1016/j.eng.2021.03.017.
Lambin, Eric F., et Helmut J. Geist. 2008. *Land-Use and Land-Cover Change: Local Processes and Global Impacts*. Springer Science & Business Media.
Lanrewaju, Funmilayo. 2012. « Urbanization, Housing Quality and Environmental Degeneration in Nigeria ». *Journal of Geography and Regional Planning* 5 (16): 422-29. https://doi.org/10.5897/JGRP12.060.
Lawson, Andrew B. 2013. *Statistical Methods in Spatial Epidemiology*. John Wiley & Sons.
Lê, Sébastien, Julie Josse, et François Husson. 2008. « **FactoMineR** : An *R* Package for Multivariate Analysis ». *Journal of Statistical Software* 25 (1). https://doi.org/10.18637/jss.v025.i01.
LECLERC, A. 1990. « Dictionnair d'epidemiologie ». *Paris : Frison-Roche* 1: 143.
Levin, Simon A. 1992. « The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture ». *Ecology* 73 (6): 1943-67. https://doi.org/10.2307/1941447.
Li, G., et Q. Weng. 2007. « Measuring the Quality of Life in City of Indianapolis by Integration of Remote Sensing and Census Data ». *International Journal of Remote Sensing* 28 (2): 249-67. https://doi.org/10.1080/01431160600735624.
Li, Huling, Hui Li, Zhongxing Ding, Zhibin Hu, Feng Chen, Kai Wang, Zhihang Peng, et Hongbing Shen. 2020. « Spatial Statistical Analysis of Coronavirus Disease 2019 (Covid-19) in China ». *Geospatial Health* 15 (1). https://doi.org/10.4081/gh.2020.867.
Linard, Catherine, et Andrew J. Tatem. 2012. « Large-scale spatial population databases in infectious disease research ». *International Journal of Health Geographics* 11 (1): 7. https://doi.org/10.1186/1476-072X-11-7.
Linard, Catherine, Andrew J. Tatem, et Marius Gilbert. 2013. « Modelling spatial patterns of urban growth in Africa ». *Applied Geography* 44 (octobre): 23-32. https://doi.org/10.1016/j.apgeog.2013.07.009.
Liu, Danning. 2014. « Determinants of Infant Mortality in Kenya - Analysis of Kenya DHS 2003 and 2008/9 ». http://lup.lub.lu.se/student-papers/record/4647031.
Lo, C. P., et Benjamin J. Faber. 1997. « Integration of landsat thematic mapper and census data for quality of life assessment ». *Remote Sensing of Environment* 62 (2): 143-57. https://doi.org/10.1016/S0034-4257(97)00088-6.

Lone, Shabir Ahmad, et Aijaz Ahmad. 2020. « COVID-19 pandemic – an African perspective ». *Emerging Microbes & Infections* 9 (1): 1300-1308. https://doi.org/10.1080/22221751.2020.1775132.

Lucas, Robyn M, et Anthony J McMichael. 2005. « Association or Causation: Evaluating Links between "Environment and Disease" ». *Bulletin of the World Health Organization*, 4.

Machault, Vanessa, Cécile Vignolles, Frédéric Pagès, Libasse Gadiaga, Abdoulaye Gaye, Cheikh Sokhna, Jean-François Trape, Jean-Pierre Lacaux, et Christophe Rogier. 2010. « Spatial heterogeneity and temporal evolution of malaria transmission risk in Dakar, Senegal, according to remotely sensed environmental data ». *Malaria Journal* 9 (1): 252. https://doi.org/10.1186/1475-2875-9-252.

« Mapping changes in housing in sub-Saharan Africa from 2000 to 2015 | Nature ». s. d. Consulté le 19 octobre 2019. https://www.nature.com/articles/s41586-019-1050-5.

Masquelier, Bruno, Lucia Hug, David Sharrow, Danzhen You, Daniel Hogan, Kenneth Hill, Jing Liu, Jon Pedersen, et Leontine Alkema. 2018. « Global, Regional, and National Mortality Trends in Older Children and Young Adolescents (5–14 Years) from 1990 to 2016: An Analysis of Empirical Data ». *The Lancet Global Health* 6 (10): e1087-99. https://doi.org/10.1016/S2214-109X(18)30353-X.

Masquelier, Bruno, Lucia Hug, David Sharrow, Danzhen You, Colin Mathers, Patrick Gerland, et Leontine Alkema. 2021. « Global, Regional, and National Mortality Trends in Youth Aged 15–24 Years between 1990 and 2019: A Systematic Analysis ». *The Lancet Global Health* 9 (4): e409-17. https://doi.org/10.1016/S2214-109X(21)00023-1.

Masquelier, Bruno, Gilles Pison, Julio Rakotonirina, et Anjarasoa Rasoanomenjanahary. 2019. « Estimating Cause-Specific Mortality in Madagascar: An Evaluation of Death Notification Data from the Capital City ». *Population Health Metrics* 17 (1): 8. https://doi.org/10.1186/s12963-019-0190-z.

Mboussou, F., P. Ndumbi, R. Ngom, Z. Kamassali, O. Ogundiran, J. Van Beek, G. Williams, C. Okot, E. L. Hamblion, et B. Impouma. 2019. « Infectious disease outbreaks in the African region: overview of events reported to the World Health Organization in 2018 ». *Epidemiology and Infection* 147 (novembre): e299. https://doi.org/10.1017/S0950268819001912.

Mollalo, Abolfazl, Behzad Vahedi, et Kiara M. Rivera. 2020. « GIS-Based Spatial Modeling of COVID-19 Incidence Rate in the Continental United States ». *Science of The Total Environment* 728 (août): 138884. https://doi.org/10.1016/j.scitotenv.2020.138884.

Moraga, Paula. 2019. *Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny*. Boca Raton: Chapman and Hall/CRC. https://doi.org/10.1201/9780429341823.

Moran, P. A. P. 1950. « Notes on Continuous Stochastic Phenomena ». *Biometrika* 37 (1/2): 17-23. https://doi.org/10.2307/2332142.

Mosley, W. Henry, et Lincoln C. Chen. 1984. « An Analytical Framework for the Study of Child Survival in Developing Countries ». *Population and Development Review* 10: 25. https://doi.org/10.2307/2807954.

Ndiaye, Ibrahima. 2015. « Étalement urbain et différenciation sociospatiale à Dakar (Sénégal) ». *Cahiers de géographie du Québec* 59 (166): 47-69. https://doi.org/10.7202/1034348ar.

Ndiaye, K, C Mathiot, A Spiegel, M Sagna, et A Dosseh. 2000. « Epidemiological and Virological Influenza Survey in Dakar, Senegal: 1996-1998. » *The American Journal of Tropical Medicine and Hygiene* 62 (5): 639-43. https://doi.org/10.4269/ajtmh.2000.62.639.
Ndonky, Alphousseyni, Sébastien Oliveau, Richard Lalou, et Stéphanie Dos Santos. 2015. « Mesure de l'accessibilité géographique aux structures de santé dans l'agglomération de Dakar ». *Cybergeo : European Journal of Geography*, décembre. https://doi.org/10.4000/cybergeo.27312.
Nguimkeu, Pierre, et Sosson Tadadjeu. 2021. « Why Is the Number of COVID-19 Cases Lower than Expected in Sub-Saharan Africa? A Cross-Sectional Analysis of the Role of Demographic and Geographic Factors ». *World Development* 138 (février): 105251. https://doi.org/10.1016/j.worlddev.2020.105251.
« Noncommunicable Diseases ». s. d. WHO | Regional Office for Africa. Consulté le 13 octobre 2021. https://www.afro.who.int/health-topics/noncommunicable-diseases.
« Noncommunicable Diseases Increase Risk of Dying from COVID-19 in Africa - World ». s. d. ReliefWeb. Consulté le 13 octobre 2021. https://reliefweb.int/report/world/noncommunicable-diseases-increase-risk-dying-covid-19-africa.
Ogoina, Dimie, et Geofrey C. Onyemelukwe. 2009. « The Role of Infections in the Emergence of Non-Communicable Diseases (NCDs): Compelling Needs for Novel Strategies in the Developing World ». *Journal of Infection and Public Health* 2 (1): 14-29. https://doi.org/10.1016/j.jiph.2009.02.001.
Omran, Abdel R. 2005. « The Epidemiologic Transition: A Theory of the Epidemiology of Population Change ». *The Milbank Quarterly* 83 (4): 731-57. https://doi.org/10.1111/j.1468-0009.2005.00398.x.
Owens, Ann. 2012. « Neighborhoods on the Rise: A Typology of Neighborhoods Experiencing Socioeconomic Ascent ». *City & Community* 11 (4): 345-69. https://doi.org/10.1111/j.1540-6040.2012.01412.x.
Oztig, L. I., et O. E. Askin. 2020. « Human Mobility and Coronavirus Disease 2019 (COVID-19): A Negative Binomial Regression Analysis ». *Public Health* 185 (août): 364-67. https://doi.org/10.1016/j.puhe.2020.07.002.
Pongou, Roland, Majid Ezzati, et Joshua A. Salomon. 2006. « Household and Community Socioeconomic and Environmental Determinants of Child Nutritional Status in Cameroon ». *BMC Public Health* 6 (1): 98. https://doi.org/10.1186/1471-2458-6-98.
Qi, C., Y. C. Zhu, C. Y. Li, Y. C. Hu, L. L. Liu, D. D. Zhang, X. Wang, et al. 2020. « Epidemiological Characteristics and Spatial–temporal Analysis of COVID-19 in Shandong Province, China ». *Epidemiology & Infection* 148. https://doi.org/10.1017/S095026882000151X.
Quentin, Wilm, Olayinka Abosede, Joseph Aka, Patricia Akweongo, Kouassi Dinard, Alex Ezeh, Ramadan Hamed, et al. 2014. « Inequalities in child mortality in ten major African cities ». *BMC Medicine* 12 (1): 95. https://doi.org/10.1186/1741-7015-12-95.
Rahman, Atiqur, Yogesh Kumar, Shahab Fazal, et Sunil Bhaskaran. 2011. « Urbanization and Quality of Urban Environment Using Remote Sensing and GIS Techniques in East Delhi-India ». *Journal of Geographic Information System* 03 (01): 62. https://doi.org/10.4236/jgis.2011.31005.

Raihan Sharif, A. H. M., Sheikh Md. Monzurul Huq, et Mesbah-us-Saleheen. 1993. « Spatial Patterns of Mortality in Bangladesh ». *Social Science & Medicine*, Special Issue Geographical Equalities of Mortality in Developing Countries, 36 (10): 1325-30. https://doi.org/10.1016/0277-9536(93)90223-Q.

« Rapport définitif RGPHAE-2013 ». 2013, 418.

Ridout, Martin, Clarice G B Demetrio, et John Hinde. s. d. « Models for Count Data with Many Zeros », 13.

Sagui, Emmanuel, Papa Saliou M'Baye, Christophe Dubecq, Khadi Ba Fall, Abdourahmane Niang, Sarah Gning, Jean-Pierre Bellefleur, Mouhamadou Sane, et Jean Marc Debonne. 2005. « Ischemic and Hemorrhagic Strokes in Dakar, Senegal ». *Stroke* 36 (9): 1844-47. https://doi.org/10.1161/01.STR.0000177864.08516.47.

Saikia, Nandita, et F. Ram. 2010. « Determinants of Adult Mortality in India ». *Asian Population Studies* 6 (2): 153-71. https://doi.org/10.1080/17441730.2010.494441.

Salem, Gérard. 1998a. *La santé dans la ville: géographie d'un petit espace dense : Pikine (Sénégal)*. KARTHALA Editions.

———. 1998b. *La santé dans la ville: géographie d'un petit espace dense : Pikine (Sénégal)*. KARTHALA Editions.

Santos, Stéphanie Dos, et Sabine Henry. 2008. « Rainfall Variation as a Factor in Child Survival in Rural Burkina Faso: The Benefit of an Event-History Analysis ». *Population, Space and Place* 14 (1): 1-20. https://doi.org/10.1002/psp.470.

Sanyaolu, Adekunle, Chuku Okorie, Aleksandra Marinkovic, Risha Patidar, Kokab Younis, Priyank Desai, Zaheeda Hosein, Inderbir Padda, Jasmine Mangat, et Mohsin Altaf. 2020. « Comorbidity and Its Impact on Patients with COVID-19 ». *SN Comprehensive Clinical Medicine* 2 (8): 1069-76. https://doi.org/10.1007/s42399-020-00363-4.

Satterthwaite, David, Alice Sverdlik, et Donald Brown. 2019. « Revealing and Responding to Multiple Health Risks in Informal Settlements in Sub-Saharan African Cities ». *Journal of Urban Health* 96 (1): 112-22. https://doi.org/10.1007/s11524-018-0264-4.

« Sénégal: Profil Urbain De Dakar (French) ». s. d. Consulté le 12 novembre 2018. https://unhabitat.org/books/senegal-profil-urbain-de-dakar-french/.

Shrestha, Nistha, Muhammad Yousaf Shad, Osman Ulvi, Modasser Hossain Khan, Ajlina Karamehic-Muratovic, Uyen-Sa D. T. Nguyen, Mahdi Baghbanzadeh, et al. 2020. « The Impact of COVID-19 on Globalization ». *One Health* 11 (décembre): 100180. https://doi.org/10.1016/j.onehlt.2020.100180.

Singh, Akansha, et Bruno Masquelier. 2018. « Continuities and changes in spatial patterns of under-five mortality at the district level in India (1991–2011) ». *International Journal of Health Geographics* 17 (1): 39. https://doi.org/10.1186/s12942-018-0159-3.

Smew, Awad I., Cecilia Lundholm, Lars Sävendahl, Paul Lichtenstein, et Catarina Almqvist. 2020. « Familial Coaggregation of Asthma and Type 1 Diabetes in Children ». *JAMA Network Open* 3 (3): e200834. https://doi.org/10.1001/jamanetworkopen.2020.0834.

Soura, Bassiahi. 2009. *Analyse de la mortalité et de la santé des enfants à Ouagadougou: Inégalités spatiales, effets individuels et effets contextuels*. Presses univ. de Louvain.

Souris, Marc. 2019. *Épidémiologie et géographie: Principes, méthodes et outils de l'analyse spatiale*. ISTE Group.

Stiegler, Nancy, et Jean-Pierre Bouchard. 2020. « South Africa: Challenges and Successes of the COVID-19 Lockdown ». *Annales Médico-Psychologiques, Revue Psychiatrique* 178 (7): 695-98. https://doi.org/10.1016/j.amp.2020.05.006.

Suglia, Shakira Franco, Cristiane S. Duarte, et Megan T. Sandel. 2011. « Housing Quality, Housing Instability, and Maternal Mental Health ». *Journal of Urban Health* 88 (6): 1105-16. https://doi.org/10.1007/s11524-011-9587-0.

Sun, Feinuo, Stephen A. Matthews, Tse-Chuan Yang, et Ming-Hsiao Hu. 2020. « A Spatial Analysis of the COVID-19 Period Prevalence in U.S. Counties through June 28, 2020: Where Geography Matters? » *Annals of Epidemiology* 52 (décembre): 54-59.e1. https://doi.org/10.1016/j.annepidem.2020.07.014.

Tais Grippa, et Stefanos Georganos. 2018. « Dakar very-high resolution land cover map ». Zenodo. https://doi.org/10.5281/zenodo.1290800.

Tatem, Andrew J., Abdisalan M. Noor, Craig von Hagen, Antonio Di Gregorio, et Simon I. Hay. 2007. « High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa ». *PLOS ONE* 2 (12): e1298. https://doi.org/10.1371/journal.pone.0001298.

Thomas, Richard K. 2018. *Concepts, Methods and Practical Applications in Applied Demography: An Introductory Textbook*. Springer International Publishing. https://doi.org/10.1007/978-3-319-65439-3.

Thomson, Curtis N, et Perry Hardin. 2000. « Remote Sensing/GIS Integration to Identify Potential Low-Income Housing Sites ». *Cities* 17 (2): 97-109. https://doi.org/10.1016/S0264-2751(00)00005-6.

Thomson, Dana R., Radheshyam Bhattarai, Sudeepa Khanal, Shraddha Manandhar, Rajeev Dhungel, Subash Gajurel, Joseph Paul Hicks, et al. 2021. « Addressing Unintentional Exclusion of Vulnerable and Mobile Households in Traditional Surveys in Kathmandu, Dhaka, and Hanoi: A Mixed-Methods Feasibility Study ». *Journal of Urban Health* 98 (1): 111-29. https://doi.org/10.1007/s11524-020-00485-z.

Trape, Jean-Francois, Evelyne Lefebvre-Zante, Fabrice Legros, Gora Ndiaye, Hilaire Bouganali, Pierre Druilhe, et Gerard Salem. 1992. « Vector Density Gradients and the Epidemiology of Urban Malaria in Dakar, Senegal ». *The American Journal of Tropical Medicine and Hygiene* 47 (2): 181-89. https://doi.org/10.4269/ajtmh.1992.47.181.

Tuson, Matthew, Matthew Yap, Mei Ruu Kok, Bryan Boruff, Kevin Murray, Alistair Vickery, Berwin A. Turlach, et David Whyatt. 2020. « Overcoming inefficiencies arising due to the impact of the modifiable areal unit problem on single-aggregation disease maps ». *International Journal of Health Geographics* 19 (1): 40. https://doi.org/10.1186/s12942-020-00236-y.

Tusting, Lucy S., Christian Bottomley, Harry Gibson, Immo Kleinschmidt, Andrew J. Tatem, Steve W. Lindsay, et Peter W. Gething. 2017. « Housing Improvements and Malaria Risk in Sub-Saharan Africa: A Multi-Country Analysis of Survey Data ». *PLOS Medicine* 14 (2): e1002234. https://doi.org/10.1371/journal.pmed.1002234.

United Nations. 2020. *World Social Report 2020: Inequality in a Rapidly Changing World*. UN. https://doi.org/10.18356/7f5d0efc-en.

« United Nations Population Division | Department of Economic and Social Affairs ». s. d. Consulté le 13 janvier 2022. https://www.un.org/en/development/desa/population/commission/sessions/2019/ngo/index.asp.

« UNSD — Welcome to UNSD ». s. d. Consulté le 13 janvier 2022. https://unstats.un.org/home/.
Utzinger, J., et J. Keiser. 2006. « Urbanization and Tropical Health — Then and Now ». *Annals of Tropical Medicine & Parasitology* 100 (5-6): 517-33. https://doi.org/10.1179/136485906X97372.
Vaid, Uchita, et Gary W. Evans. 2017. « Housing Quality and Health: An Evaluation of Slum Rehabilitation in India ». *Environment and Behavior* 49 (7): 771-90. https://doi.org/10.1177/0013916516667975.
Wagstaff, Adam. 2002. « Poverty and Health Sector Inequalities ». *Bulletin of the World Health Organization* 80: 97-105. https://doi.org/10.1590/S0042-96862002000200004.
Wang, Haidong, Amanuel Alemu Abajobir, Kalkidan Hassen Abate, Cristiana Abbafati, Kaja M Abbas, Foad Abd-Allah, Semaw Ferede Abera, et al. 2017. « Global, Regional, and National under-5 Mortality, Adult Mortality, Age-Specific Mortality, and Life Expectancy, 1970–2016: A Systematic Analysis for the Global Burden of Disease Study 2016 ». *The Lancet* 390 (10100): 1084-1150. https://doi.org/10.1016/S0140-6736(17)31833-0.
Wen, Frank, Dongmei Chen, et Anna Majury. 2014. « Spatial Analysis and Statistical Modeling of 2009 H1N1 Pandemic in the Greater Toronto Area ». In *Analyzing and Modeling Spatial and Temporal Dynamics of Infectious Diseases*, 247-62. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118630013.ch13.
Whitehead, H. 1869. « Remarks on the Outbreak of Cholera in Broad Street, Golden Square, London, in 1854 ». *Transactions. Epidemiological Society of London* 3 (Pt 1): 99-104.
World Health Organization. 2017. *World Malaria Report 2017*. Geneva: World Health Organization. https://apps.who.int/iris/handle/10665/259492.
« World Health Statistics ». s. d. Consulté le 21 octobre 2021. https://www.who.int/data/gho/data/themes/world-health-statistics.
« World Population Prospects - Population Division - United Nations ». s. d. Consulté le 28 septembre 2020. https://population.un.org/wpp/.
« World Urbanization Prospects - Population Division - United Nations ». s. d. Consulté le 13 novembre 2018. https://population.un.org/wup/.
Yang, Tse-Chuan, Carla Shoff, et Stephen A. Matthews. 2013. « Examining the Spatially Non- Stationary Associations Between the Second Demographic Transition and Infant Mortality: A Poisson GWR Approach ». *Spatial Demography* 1 (1): 17-40. https://doi.org/10.1007/BF03354885.
Yaya, Sanni, Ghose Bishwajit, Friday Okonofua, et Olalekan A. Uthman. 2018. « Under Five Mortality Patterns and Associated Maternal Risk Factors in Sub-Saharan Africa: A Multi-Country Analysis ». *PLOS ONE* 13 (10): e0205977. https://doi.org/10.1371/journal.pone.0205977.
Yaya, Sanni, Alzahra Hudani, Ogochukwu Udenigwe, Vaibhav Shah, Michael Ekholuenetale, et Ghose Bishwajit. 2018. « Improving Water, Sanitation and Hygiene Practices, and Housing Quality to Prevent Diarrhea among Under-Five Children in Nigeria ». *Tropical Medicine and Infectious Disease* 3 (2): 41. https://doi.org/10.3390/tropicalmed3020041.
Yaya, Sanni, Olalekan A. Uthman, Friday Okonofua, et Ghose Bishwajit. 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 (1): 616. https://doi.org/10.1186/s12889-019-6940-9.

Yin, Ling, Hao Zhang, Yuan Li, Kang Liu, Tianmu Chen, Wei Luo, Shengjie Lai, et al. 2021. « A Data Driven Agent-Based Model That Recommends Non-Pharmaceutical Interventions to Suppress Coronavirus Disease 2019 Resurgence in Megacities ». *Journal of The Royal Society Interface* 18 (181): 20210112. https://doi.org/10.1098/rsif.2021.0112.

Zhou, Chenghu, Fenzhen Su, Tao Pei, An Zhang, Yunyan Du, Bin Luo, Zhidong Cao, et al. 2020. « COVID-19: Challenges to GIS with Big Data ». *Geography and Sustainability* 1 (1): 77-87. https://doi.org/10.1016/j.geosus.2020.03.005.

List of figures

List of tables