University of Southampton

Faculty of Environmental and Life Sciences

School of Geography and Environmental Science

Understanding Urban Water Utility Performance, Water Network Topology and Geographic Water Access: Case Studies in Sub-Sahara Africa

by

Zhangliang Deng

ORCID ID: https://orcid.org/0000-0002-0080-0564

Thesis for the degree of Doctor of Philosophy

October 2025

University of Southampton

Abstract

Faculty of Environmental and Life Sciences

School of Geography and Environmental Science

Doctor of Philosophy

Understanding Urban Water Utility Performance, Water Network Topology and Geographic Water Access: Case Studies in Sub-Sahara Africa

by

Zhangliang Deng

Since the United Nations launched initiatives to improve global water governance and resilience, progress has been made towards Sustainable Development Goal (SDG) 6, which aims to ensure universal access to safe water. However, significant challenges remain, particularly concerning inequalities in urban water supply, a problem that is especially pronounced in Sub-Saharan Africa (SSA). These disparities are closely linked to the region's urbanisation processes, often characterised by the proliferation of slums, which SDG 11 (sustainable and safe cities) aims to upgrade. While numerous studies have examined the heterogeneity of urban water services in SSA, there has been a lack of quantitative analysis on how urban environments influence water service performance. Focusing on water pipeline networks and kiosk-based services in two SSA cities, 1) Kisumu, Kenya, and 2) Kigali, Rwanda, the research integrates graph-based modelling, co-location analysis, network-based community detection algorithms, and the two-step floating catchment area (2SFCA) model to assess the impact of urban morphology, slum distribution, and policy interventions on both piped and non-piped water services. The findings from the graph-based analysis indicate that the configuration of essential services, particularly piped water provision, is shaped by two primary factors: 1) The principles guiding network construction, and 2) The influence of urban morphology. Pipeline networks exhibit discernible correlations with road networks, with betweenness and closeness centralities displaying similar distributions across both cities. Community detection further reveals that pipelines serving slum areas form distinct clusters from those supplying other neighbourhoods. In Kisumu, areas under delegated water service management arrangements also form distinct networks. Accessibility analysis of water kiosks highlights disparities in water point availability within slums, shaped in part by broader urban layout constraints. Additionally, the study finds that the methodology used to generate population datasets significantly influences water access indices. Settlement-constrained datasets offer a more robust representation of water access in SSA cities. Given the limited prior applications of graph-based and 2SFCA methods in water service research, this study provides a quantitative workflow for assessing urban water service disparities in data-scarce SSA cities and addresses critical knowledge gaps in both water governance and urban studies. The findings of this study underscore the need for an integrated approach, in which water provision challenges should be addressed alongside broader urban planning initiatives such as slum upgrading programmes.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Zhangliang Deng

Title of thesis: Understanding Urban Water Utility Performance, Water Network Topology and Geographic Water Access: Case Studies in Sub-Sahara Africa

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
   The main content of Chapter 5 has been presented as a poster at the IWA Efficient Urban Water Management Conference (2023, Bordeaux, France).
   (Accepted) ‘Graph and Community Detection Analysis of Pipeline Network Configuration and Urban Morphology in Kisumu and Kigali’ [submitted to Journal of Water Resources Planning and Management]

Signature: .... Date:........................

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my main supervisor, Professor James Wright. He has been my mentor since my master's degree, a period that coincided with the challenges of the COVID-19 pandemic. Without his unwavering support, I would never have been able to pursue my academic journey so deeply. Throughout my entire PhD, he devoted countless hours and immense care to guiding me in all aspects. From inspiring new ideas to helping with writing, and even the small but significant details of the PhD life, this thesis would not have come to fruition without his dedicated mentorship.

I am also sincerely grateful to my second supervisor, Dr Christopher Lloyd. His enthusiasm and energy brought a fresh and vital perspective to this PhD project. With his technical expertise, he provided detailed and insightful guidance on the technical aspects of the research, helping me to think more rigorously and consider important technical nuances.

My thanks also go to Professor David Martin, whose wisdom and profound academic insight have been an invaluable source of inspiration. From the very beginning of my research, he offered many guiding suggestions and was always approachable and generous with his time. His timely responses to my queries and continuous support throughout my doctoral studies have left a lasting impression on me. I sincerely wish him good health and happiness in his retirement.

I would also like to thank Lorna-Grace Okotto and Joseph Okotto-Okotto for generously providing the data essential to this research. Their deep local knowledge and insightful advice have greatly enriched this study.

Finally, I would like to thank the university community for providing a supportive academic environment and valuable assistance throughout my studies.

Most importantly, I am deeply grateful to my family and friends for their unwavering encouragement and support during this journey.

Chapter 1 Introduction

In 2015, the world's development agenda shifted from the Millennium Development Goals (MDGs) to the Sustainable Development Goals (SDGs). This 15-year agenda encompasses various aspects such as health, hunger, water, and urban and community development, with widespread inequality in cities identified as a major issue to be addressed (UN-habitat, 2016). However, a decade after the launch of the SDGs, the achievement of many targets remains uncertain, with significant disparities persisting across regions, particularly in the area of water services. According to the United Nations' 2023 report, 703 million people still lacked access to basic water services, of whom 408 million were located in sub-Saharan Africa (SSA) (UN, 2023a). The causes of water scarcity extend beyond shifts in lifestyle and consumption patterns; a key driver has been population growth, which has led to demand outstripping supply, while uneven distribution further exacerbates the uncertainty of water access (Molden, 2020, Mukheibir, 2010). Schlosser et al. (2014) estimated that, by 2050, economic and demographic factors would cause an additional 1.8 billion people to suffer from water scarcity, and most of these populations will be located in developing countries.

The disadvantaged position of Southern cities in terms of water supply is closely linked to their urbanisation processes. The rapid urbanisation that has characterised the twenty-first century has generated an insatiable demand for housing and basic services (UN-Habitat, 2023a). Meanwhile, rapid urbanisation has widened disparities in service quality within cities (Boakye-Ansah et al., 2019, Ocholla et al., 2022, UN-ECA, 2014). In the context of rapid urbanisation in Southern cities, these challenges have contributed to the emergence of slums—densely populated, underserved communities that embody the imbalance between supply and demand for urban resources and services. Consequently, water access issues and the development of slums are closely intertwined, both spatially and conceptually.

SDG 11 ("Make cities and human settlements inclusive, safe, resilient, and sustainable") is closely connected to the objectives of SDG 6 ("Ensure availability and sustainable management of water and sanitation for all") regarding equitable water provision. The close linkage between these two goals is reflected in policy practices, such as the emphasis on infrastructure services, particularly water services, within slum upgrading programmes (Adama, 2020, Brown-Luthango et al., 2017, Olthuis et al., 2015), and water interventions specifically targeting slums (Annamalai et al., 2016, Dos Santos et al., 2017, Lima et al., 2021, Marin, 2009b, Moretto et al., 2018, UN-Habitat, 2023b). Theoretical frameworks have also explored the connections between infrastructure and urbanisation, including discussions of the splintering of infrastructure networks during urban growth (Coutard, 2008, Graham and Marvin, 2002), the theorisation of how Southern urban characteristics influence governance and modes of production (Parida and Agrawal, 2023), and research linking urban form to network performance(Lorenz et al., 2021, Torres et al., 2017, Zhao et al., 2020).

Nonetheless, there remains a notable evidence gap regarding the current status of water services within Southern cities and their synergies with slum characteristics and policies. Existing research on urban infrastructure has predominantly focused on the relationship between road networks and urban morphology, with graph theory methods widely applied to reveal the connections between urban networks and morphological or environmental factors (Dingil et al., 2018, Dovey et al., 2020, Kolowa et al., 2024, Serra et al., 2016, Yin et al., 2018). By contrast, studies concerning water services are significantly lacking, and those that do exist are largely concentrated in Northern cities. Little is known about how urban water supply services—including both piped networks and water points such as kiosks—are influenced by urban form. This is particularly pertinent in the context of Southern cities, where slums form an integral part of the urban landscape and may shape the spatial configuration of water infrastructure. Furthermore, Southern cities have implemented various policies to intervene in water delivery within slums (Adams et al., 2019, Dos Santos et al., 2017, Patel and Killemsetty, 2020, World Bank, 2009), potentially adding another layer of complexity. However, there is limited understanding of the extent to which these policies have physically influenced water infrastructure.

Therefore, this research aims to analyse the relationship between the SSA urban environment and the performance of water services. It is necessary to examine the spatial patterns of piped infrastructure in selected case study areas to inform the factors influencing urban water service delivery, particularly in slums of larger cities in SSA. Specific objectives include:

1. Assess the extent of topological and geometrical commonalities of road and water/wastewater networks across case study SSA cities and evaluate to what extent road typologies capture the distribution of slums and the heterogeneity of water and sanitation service infrastructures.
2. Assess whether the spatial distribution and topology of water networks differ between slum versus planned urban areas and between water management regimes.
3. Quantify geographic access to kiosk water in a case study city by utilising the 2 steps floating catchment area (2SFCA) model and analyse the extent to which access to kiosk water points is influenced by urban planning.

Against this backdrop, Chapter 2 provides a review of the progress of urbanisation and SDG initiatives in Southern cities, as well as the policies concerning slums and water supply that are prominent in Southern urban contexts. The review also introduces the two-step floating catchment area (2SFCA) method, graph theory metrics and algorithms, and relevant application contexts. A detailed summary of the study areas' urbanisation patterns and infrastructure development is presented in Chapter 3. The analytical chapters are organised to address these research objectives from various perspectives (Figure 1.1). The first chapter (Chapter 4) quantifies co-location and topological similarities in road, water, and wastewater urban networks in Kigali versus Kisumu, summarising patterns found. The second chapter (Chapter 5) examines the pipeline network's relationship with slum distribution and policy by applying graph analysis and community detection to Kisumu and Kigali's pipeline networks. The third (Chapter 6) focuses on spatial patterns of water service accessibility by examining another major water source in the southern cities: kiosks. Due to data availability, 2SFCA results are presented for Kisumu's kiosks only.

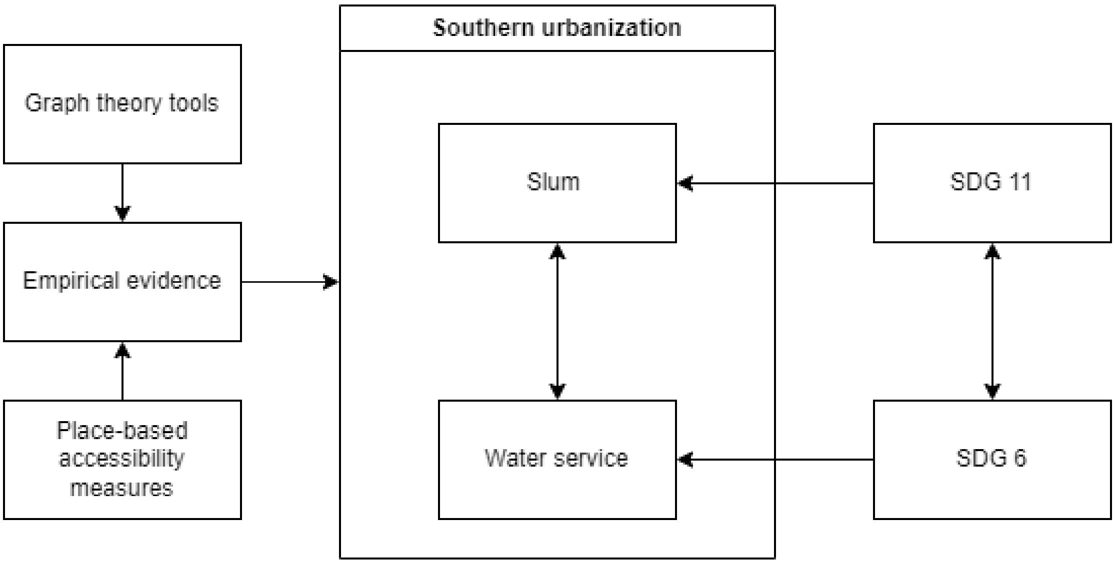


Figure 1.1 Conceptual Framework of the Thesis. In the context of southern urbanisation, there are notable similarities between the trajectories of slum improvement and water policy development. This research aims to identify the connections between these two aspects, providing empirical evidence for the study of southern cities and ultimately contributing to the understanding and achievement of SDG 6 and SDG 11.

Chapter 2 Literature review

2.1 The Sustainable Development Goals (SDGs)

The Sustainable Development Goals (SDGs) were launched by the United Nations in 2015 as a global initiative aimed at eradicating poverty and promoting sustainable development by 2030. The SDGs include 17 goals, 169 targets, and over 230 indicators, addressing issues such as poverty, health, environment, energy, and urban development, which can be classified into three themes: economic growth, environmental protection, and social inclusion (Feeny, 2020, Herrera, 2019). According to the 2030 agenda (UN, 2015a), the emphasis is on both the differences between countries and the disparities within countries. In other words, the SDGs aim to leave no one behind on the road to development, whether in developed or developing countries. This is a historic shift, implying not only the need for multidimensional participation in governance (participation of organisations from local to global level) but also a greater focus on addressing inequality and poverty within countries (Revi, 2016). Such efforts are closely linked to urbanisation (UN-habitat, 2016).

Today, some 56% of the world's population lives in cities, and more than 80% of global GDP is generated in cities (World Bank, 2023). Since before 1960, the total global population has risen, and so has the proportion of people living in cities. Until relatively recently this trend towards urbanisation has seemed irreversible (Chen et al., 2014). Historically, when economies of scale develop within a city, the growing population moves into the city to participate in productive activities, thus contributing to urbanisation. Therefore, the population growth accompanying urbanisation is conducive to economic development - this can be seen from the share of urban GDP. It is also argued that urbanisation will facilitate the flow of goods and services, stimulate the rural economy and narrow the gap between urban and rural areas, thereby reducing poverty and inequality (Ahimah-Agyakwah et al., 2022).

The previously described circumstance is not always the case. As noted by Gollin et al. (2016), urbanisation in Western cities has typically been accompanied by a transformation in economic structure, driven by the concentration of population in urban industrial sectors. In contrast, the trajectory of urbanisation in Sub-Saharan Africa (SSA) has diverged from trends in the share of the manufacturing sector in GDP (see Figure 2.1). A substantial proportion of the urban population in SSA remains employed in the agricultural sector (Grover et al., 2022). The ongoing urbanisation in SSA has been characterised as urbanisation without economic growth (Castells-Quintana and Wenban-Smith, 2020), or so-called *premature urbanisation* (Grover et al., 2022). Moreover, urbanisation in SSA is accompanied by growing inequality in cities and problems such as health risks (Alaazi and Aganah, 2020, Amegah, 2021). Thus, concerns have been raised about the living conditions of SSA urban dwellers.

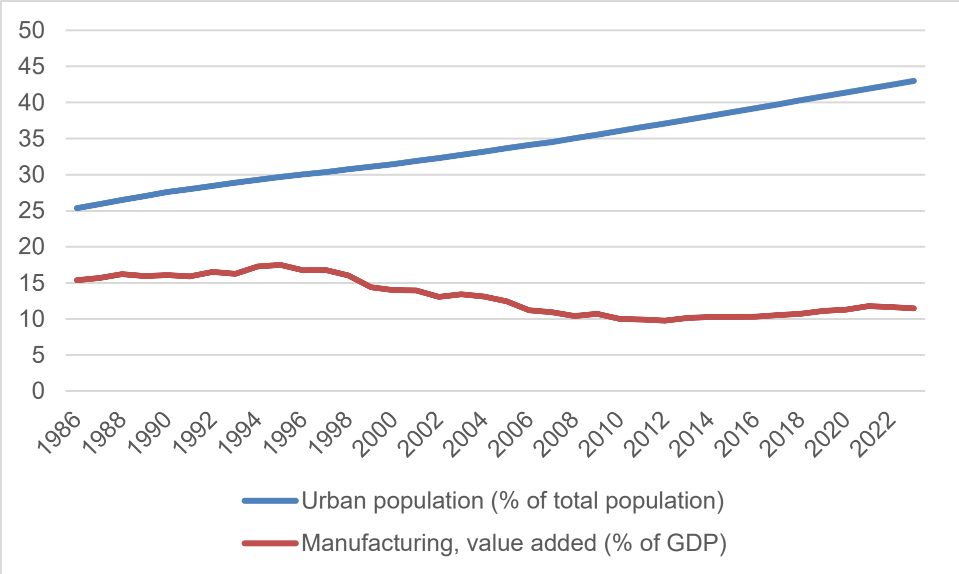


Figure 2.1 Trends in Urbanisation Rate and Manufacturing Share of GDP in SSA, 1986–2023 (Source: World Development Indicators (accessed 2025)).

2.1.1 SDGs in focus: SDG 11- sustainable cities

Based on urbanisation trends and impacts in regions such as SSA, SDG 11 was proposed to address issues in the urban development process. As shown in Table 2.1, the overarching aim of SDG 11 is to make cities and human settlements inclusive, safe, resilient and sustainable. This framework covers many aspects of urban life.

Table 2.1 Targets, and indicators for SDG 11 derived from (UN, 2015b). Italicized items indicate targets and indicators relevant to the research theme.

Goal 11. Make cities and human settlements inclusive, safe, resilient and sustainable	
Target	Indicator
11.1 Adequate, safe and affordable housing	*11.1.1 Urban population living in slums*
11.2 Accessible and Sustainable transport systems for all	11.2.1 Public transport access
11.3 Inclusive and Sustainable urbanisation	11.3.1 Sustainable urbanisation rates
	11.3.2 Urban planning management
11.4 Safeguard the world's cultural and natural heritage	11.4.1 Protecting cultural heritage
11.5 Reduce the number of people affected by disasters	11.5.1 Deaths and injuries from natural disasters
	11.5.2 Economic losses from natural disasters
	11.5.3 Damage to critical infrastructure and disruptions to basic services
11.6 Reduce the environment impact of cities	11.6.1 Solid waste management
	11.6.2 Urban air pollution
11.7 Provide Universal access to safe public spaces	11.7.1 Open spaces in cities
	11.7.2 Safe spaces in cities
11.a Support links between urban, peri-urban and rural areas	11.a.1 Urban and regional planning
11.b Increase integrated policies and plans towards mitigation and adaptation to climate change	11.b.1 Integrated disaster risk management
	11.b.2 Local disaster risk management
11.c Building sustainable and resilient buildings utilizing local materials	

SDGs set out detailed targets and indicators to monitor progress. These include providing accessible and affordable transport systems, reducing urban expansion, increasing participation in urban governance, enhancing cultural and heritage preservation, addressing urban resilience and climate change challenges, better management of urban environments (pollution and waste management), providing access to urban environment, and providing a comprehensive and integrated approach to urban development (Franco et al., 2019).

Despite having detailed target sets, SDG 11's progress worldwide is not encouraging, and research targeting urbanisation is still urgently needed. The UN-Habitat Synthesis Report 2023 (UN-Habitat, 2023a) indicates that the world is still far from achieving SDG11, particularly regarding the issue of slum expansion (for the definition and description of this issue, see section 2.2). Although the proportion of urban population living in slums has decreased, the absolute number has grown, currently standing at 1.1 billion (UN, 2023b). Among the regions with the highest concentration of slum populations—Central and Southern Asia (359 million), Eastern and South-Eastern Asia (306 million), and SSA (230 million)— have the highest proportion of urban population living in slums (UN-Habitat, 2023c). Additionally, the rate of reduction in both the slum population and proportion in SSA is relatively slow (UN-Habitat, 2023b), making these countries further from the goal of reducing the urban population living in slums compared to other areas (Halkos and Gkampoura, 2021). The reasons are multifaceted. It is argued that cities have significantly different baselines in terms of human activities, economic activities, cultural context, political factors, and ecology (UN, 2015a). Therefore, SDG 11 imposes targets that may be impossible for some cities to achieve (Croese et al., 2021, Janoušková et al., 2018). For example, research in SSA has shown that a lack of political will, limited funding, and weak management capacity have severely constrained the SDG 11 process (Juju et al., 2020). Additionally, the UN-Habitat report emphasises the role of urbanisation in driving slum expansion, which has been exacerbated by the COVID-19 pandemic (Miranda et al., 2023, UN-Habitat, 2023c). Given the approaching 2030 deadline, more effort needs to be directed towards areas such as SSA cities and slums.

2.1.2 SDG 6: progress and challenges in achieving universal water access

SDG 6 has encountered various challenges. Compared to Millenium Development Goal (MDG) 7c, which only targets access to improved drinking water sources and sanitation, the ambitions of SDG 6 are broader and more specific (Herrera, 2019). This goal considers both the disparities in individuals' ability to access water due to gender, income, and education level, and the spatial inequalities in resource distribution (Dos Santos et al., 2017). There are various targets displayed in Table 2.2: drinking water (SDG 6.1), hygiene and sanitation (SDG 6.2), water quality and wastewater (SDG 6.3), water use and scarcity (SDG 6.4), water resources management (SDG 6.5), water-related ecosystems (SDG 6.6), cooperation and capacity-building (SDG 6a), and participation of local communities (SDG 6b). SDG 6 commits to equal and universal access to safely managed water and sanitation services. A safely managed water service is an improved water source located on-premises, available when needed, and free from contamination (WHO/UNICEF, 2018). Water sources that do not meet these criteria but take less than 30 minutes to draw water are defined as basic services (WHO/UNICEF, 2018). By this definition, if a water source is unprotected, it will be classified as unimproved, and improved sources are "those that have the potential to deliver safe water by nature of their design and construction" (WHO/UNICEF, 2010). Improved sources include piped supplies, such as water piped into dwellings and compounds, or public taps and standpipes, as well as non-piped supplies, such as protected wells and springs. Additionally, technologies that distribute water, such as tanker trucks and carts with small tanks or drums, are also considered improved sources.

Table 2.2 Targets, and indicators for SDG 6 derived from (UN, 2015b). Italicized items indicate targets and indicators relevant to the research theme.

Goal 6. Ensure availability and sustainable management of water and sanitation for all	
Target	Indicator
6.1 Safe and affordable drinking water	*6.1.1 Safe drinking water*
6.2 End open defecation and provide access to sanitation and hygiene	6.2.1 Safe sanitation and hygiene
6.3 Improve water quality, wastewater treatment and safe reuse	6.3.1 Wastewater safety
	6.3.2 Ambient water quality
6.4 Increase water use efficiency and ensure freshwater supplies	6.4.1 Water use efficiency
	6.4.2 Levels of freshwater stress
6.5 Implement integrated water resources management	6.5.1 Integrated water management
	6.5.2 Transboundary water cooperation
6.6 Protect and restore water-related ecosystems	6.6.1 Protect and restore water-related ecosystems
6.a Expand water and sanitation support to developing countries	6.a.1 Water and sanitation support
6.b Support local engagement in water and sanitation management	6.b.1 Local participation in water and sanitation management

Under this new framework, as of 2020, 74.3% of the global population had access to safely managed drinking water. However, approximately 2 billion people still lacked access to safely managed drinking water, and 703 million did not have access to basic water services. Of these 703 million, 408 million lived in SSA (UN, 2023a). From the experience of the International Drinking Water Supply and Sanitation Decade (IDWSSD) and prior MDGs, the primary reason for low levels of water access is population growth or the discrepancy between supply and demand (Adams et al., 2019, Najlis and Edwards, 1991, Sambu and Tarhule, 2013, UN-ECA, 2014, WHO, 1992). As summarized by Huang et al. (2021) and Schlosser et al. (2014), the main drivers of water scarcity are increased water demand due to population growth and economic activity rather than reduced supply due to climatic factors. In urban areas, water supply systems are often complex and require more investment and technological input than in rural areas (Adams et al., 2019), which contributes to significant disparities in water accessibility across different urban populations in SSA (Armah et al., 2018). Notably, in slums where population density is high and resources are scarce, the increase in water demand often leads to a reliance on informal water supplies (Juju et al., 2020). One study observed that in two SSA cities, Dar es Salaam and Addis Ababa, decentralised, on-site infrastructures rather than conventional centralised water infrastructure constitute the main sources of water provision in informal settlements (Herslund and Mguni, 2019). Further research indicates that water insecurity in slums is a widespread issue across many SSA regions (Dos Santos et al., 2017, Nyika and Dinka, 2023). With the rapid population growth and governance failures in SSA, SDG 6 is expected to face even greater challenges in the foreseeable future (Adams et al., 2019, Nyika and Dinka, 2023).

2.1.3 Integrated sustainability: interactions of SDG 6 and SDG 11

Due to the multifaceted barriers to water access, achieving SDG 6 requires consideration of synergies with other SDGs. Ait-Kadi (2016) argues that, as the development goals are comprehensive and sustainable, the overlap and the interaction between SDGs reflect that the solution of one main goal involves the progress of other goals and requires universal efforts of various stakeholders. Therefore, actions against a certain indicator will inevitably positively or negatively impact other goals, and targeting only one weak point will have little success (Abson et al., 2017, Nilsson et al., 2016). Furthermore, the adverse effects of development policies in some sectors also influence others (Blanc et al., 2017). To provide an integrated approach to achieve the SDGs, some studies have explored the synergies between them. Le Blanc (2015) analysed goals based on their wording. In other words, a relationship exists between two SDGs when a target of one SDG refers to a term that is relevant to another SDG. Coopman et al. (2016) adopted a similar approach, but further classified the relationship as supporting, enabling/disenabling, and relying. UN-ESCAP (2017) analysed the relationship between clean water and sanitation (SDG 6) and other SDGs by examining the cause-and-effect relationships between targets (Figure 2.2). Other studies have described the interactions of the SDGs based on cases of different regions (Griggs et al., 2017, Herrera, 2019).

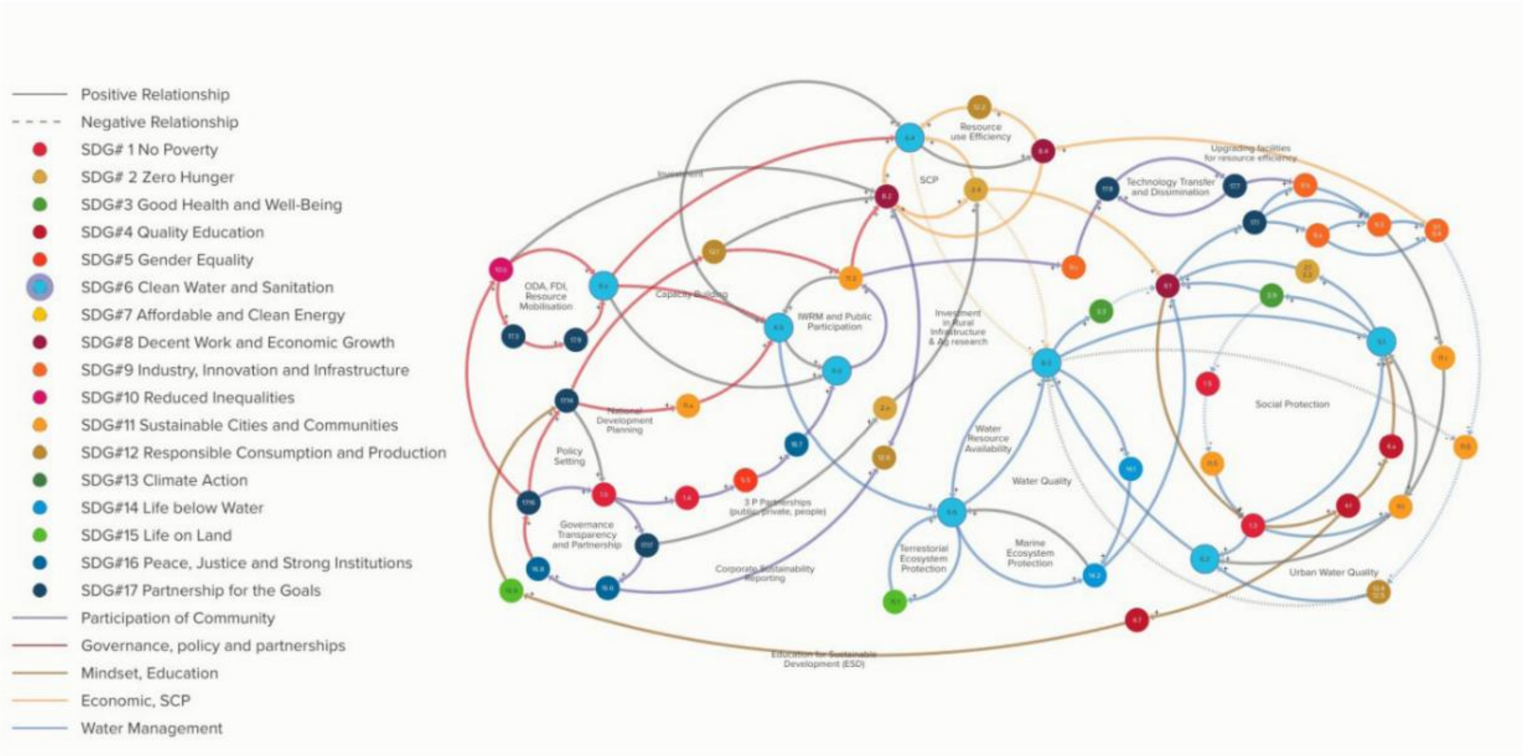


Figure 2.2 A simplified model of the interactions between the Sustainable Development Goals (UN-ESCAP, 2017).

Certain observations have been made regarding SDG 6 and SDG 11 through relevant studies. Table 2.1 shows that SDG 11 includes targets of improving infrastructure, the urban environment, environmental quality, and basic services. The "basic services" include water and waste management. Le Blanc (2015), UN-ESCAP (2017) also demonstrated that some of the objectives of SDG 6 are notably aligned with those of SDG 11 (e.g. SDG 6.1, SDG 6.3 and SDG 11.1, SDG 11.5, SDG 11.6) (see Figure 2.2). Kuc-Czarnecka et al. (2023) , based on the implementation of SDGs in European Union countries, pointed out that SDG 6 is positively correlated with almost all other goals, including SDG 11. Past United Nations practices also found a strong link between water and sanitation and urbanisation in developing countries (Adams et al., 2019, Najlis and Edwards, 1991, UN-ECA, 2014). Similarly, Dos Santos et al. (2017) point out that in SSA, the population increase is largely concentrated in urban areas and informal settlements. Therefore, holistic urban planning is crucial for meeting their water demands. A limitation of these studies is that most analyses remain at a conceptual or policy-discursive level, grounded in interpretations of SDG semantics or system-wide interactions. In addition, the nature of the interlinkages between these SDGs varies significantly depending on contextual factors (governance, technology, time, geographical location, for instance) (Blanc et al., 2017, Coopman et al., 2016, UN-Water, 2021). This highlights a lack of empirical evidence on how urbanisation processes and the development of slums concretely influence water and sanitation systems in specific regional contexts.

In summary, the SDGs form an initiative that concentrates the efforts of multidimensional participants and aims to banish poverty worldwide. SDGs focus on cities rather than countries, given the prevalence of urban inequality and poverty. Regions fall behind for many reasons, and the global indicator set is hampered in practice by many aspects. Further research is still needed to understand the underlying drivers behind lagging regions. Another innovation of the SDGs is the development of new monitoring indicators for SDG 6. Although the UN (2015a) claimed that the water target for MDG 7c was met globally in 2010, some argued that this target was built on an imperfect monitoring system (Bain et al., 2012). To better reflect the ambition of SDG6, safely managed water was adopted as a new 2030 target (WHO, 2017, WHO/UNICEF, 2015). According to its definition, many developing countries are currently far from this goal (UN, 2023b). A shared concern of both SDG 11 and SDG 6 is slums, which not only manifest the population and resource challenges of urbanisation but also hinder the expansion of water service coverage. Therefore, slum elimination is the first target of SDG 11 and one of the most critical issues in its implementation (UN-Habitat, 2023b). Simultaneously, slums mirror the challenges of SDG 6, highlighting significant spatial disparities in water access. Although there is substantial research on the synergies and trade-offs between SDG 11 and SDG 6, few studies have focused specifically on slums. Therefore, the following sections will present the practical implementations of both SDGs and their interconnections.

2.2 Slum formation, evolution and water policy

According to UN-Habitat (2022), as of 2020, over 1 billion people reside in slums and are subject to the worst forms of deprivation and marginalisation. It is estimated that by 2050, the number of people living in slums could reach two billion. The United Nations estimate regarding slum populations draws on some elements of the historical concept of “slum,” which originated in nineteenth-century London (Harris, 2009). The Industrial Revolution triggered rapid population growth and urbanisation – creating many new city spaces, including slums (Tockner and Stanford, 2002). These slums are typically characterized by overcrowding, informal housing, inadequate safe water and sanitation, and insecure tenure (land encroachment) (Davis, 2004, UN - Habitat, 2004). Beyond the United Nations, there are varying perspectives on how slums should be understood, and reviewing these different views can help deepen our understanding of the nature of slums and the current state of urban water policies.

2.2.1 Slum and urban informality: changing perspectives

With the launch of the United Nations "Cities Without Slums" initiative in 1999, the term slum has returned to prominence, and is also incorporated into the SDGs (Gilbert, 2007, UN-Habitat, 2018). Some SDGs have a focus on urban poverty and inequality: SDG 11 and SDG 6 both deal with slums. To unpack the relationship between slums and urbanisation, population growth, and basic services, it is necessary to explore what a slum is and how it is formed.

UN-Habitat (2006) defines a slum household as one or a group of individuals living under the same roof in an urban area, lacking in one or more of the following five amenities:

1. Durable housing of a permanent nature that protects against extreme climate conditions.
2. Sufficient living space, defined as not more than three people sharing the same habitable room.
3. Easy access to safe water in sufficient amounts at an affordable price.
4. Access to adequate sanitation in the form of a private or public toilet shared by a reasonable number of people.
5. Security of tenure that prevents forced evictions.

The five criteria reference a measurable baseline for the minimum materials needed for human habitation, thereby enabling clear understanding of the global urbanisation process (Lucci et al., 2018). It is an international definition that helps identify the main challenges of SDG 11. Based on data sources that include these indicators (primarily census and national household surveys), the specialized agency for sustainable urbanisation and human settlements, UN-Habitat, is able to monitor and report on slums and the other eight SDG 11 indicators (UN-Habitat, 2021). It is noteworthy that, due to difficulties in defining and assessing it, tenure security is not currently used by UN-Habitat in the measurement of slums (Dovey et al., 2021b, UN-Habitat, 2006, UN-Habitat, 2018).

Despite this intuitive definition, many other terms and standards have been adopted (Table 2.3). Areas defined as slums by the UN may be called *conventillos*, *favela* or *umjondolo*, and the definitions are also very different (Criqui, 2018, Moreno, 2003, Portes, 1971). For example, Uganda's definition of slums includes criteria such as unemployment, low income, noise, crime, drug abuse, immorality, alcoholism, HIV/ AIDS prevalence and fragile location (MLHUD, 2008). Meanwhile, slum identification in India is primarily based on legal designation (Nolan, 2015). This diversity in slum definitions partly arises from various slum patterns. The formation of slums is influenced by many factors such as geographic location, climate, politics, and culture (Kuffer et al., 2016). Therefore, slums have diverse forms: within them, housing quality may depend on locally available materials (e.g. metal, cardboard, plastic), and building density and height change as slums evolve, so newly developed slums differ from those that are more established. Consequently, it is difficult to define these settlements within one set standard (Gilbert, 2007). The varying national definitions of slums mean that slum populations are enumerated differently between countries (Kuffer et al., 2018, UN-Habitat, 2018). This lack of harmonisation is a major reason why United Nations statistics on slums are often inaccurate and show large fluctuations (Ezeh et al., 2017, Nolan, 2015), as they primarily rely on population and housing censuses and national surveys (Ezeh et al., 2017).

Table 2.3 Terms Used by Different Countries for "Informal Settlements" (Criqui, 2018)

Country	Name
Angola	Musseque
Argentina	Villa miseria
Brazil	Favela
Canada	Shantytown
Ecuador	Invasiones
France / Morocco	Bidonville
India (Delhi)	Jhuggi
Indonesia	Kampung
Madagascar	Bas quartiers
Mauritania	Gazra
Mexico	Colonias populares
Niger	Village urbain
Peru	Asentamiento humano
Philippines	Squatter settlements
South Africa	Township
Syria	Mukhalafat
Tunisia	Quartiers populaires
Turkey	Geçekondu
UK	Slum

Among varied terms and definitions, "informal settlement" deserves special mention. Informal settlement is another term used by the UN that is related to slums. Due to the long-standing stereotype associated with the word "slum" (Andavarapu and Edelman, 2013, Gilbert, 2007, Harris, 2020), informal settlement is considered a synonym for 'slum' (Dovey et al., 2021b). According to UN habitat, informal settlements are defined by three main criteria which are already covered in the definition of slums (UN-Habitat, 2021):

1. Insecurity of tenure
2. Lack of formal basic services and city infrastructure,
3. Housing not in compliance with planning and building regulations.

Many have questioned whether "informal settlement" should replace the term "slum." Here, the study does not intend to differentiate between "informal settlement" and "slum." Interested readers can refer to the discussions of Khan et al. (2023) and Dovey et al. (2021a) on this issue. Nevertheless, it is important to note that the concept of informality provides a clearer perspective on the formation of slums and the inequalities in infrastructure within them (Dovey et al., 2021b, Dovey et al., 2020). By definition, "informality" does not mean illegality, but rather activities that are not regulated by the authorities (Bastia, 2015, Charmes, 2012, Lara-Hernandez et al., 2020). After Keith Hart's research (Hart, 1973) introduced the concept of the informal economy to urban life, people began to realize that the distinction between formality and informality exists widely in cities beyond the economic sector (Chen, 2012, Recio et al., 2017). As demonstrated by Banks et al. (2020), urban informality has economic, spatial, and political dimensions that are not easily separated. Parida and Agrawal (2023) identify informality as one of the seven key characteristics of southern cities. Slums, in particular, can be seen as places where informal practices can readily be observed (Roy, 2011).

More simply, the primary drivers of slum formation worldwide are associated with the issue of supply and growing demand (Abass and Kucukmehmetoglu, 2021, Smit et al., 2017, Woo and Jun, 2020). Schindler (2017) points out that in industrialized cities, it is capital that transforms rural populations into workers. This process is driven by two forces. Typical push factors of slum formation include the decline of the agricultural economy, the degradation of land quality and quantity, or the unfavourable social environment in rural areas. Pull forces from cities include better job prospects, expectations of higher urban living standards, and risk mitigation from relying on a single rural income source (Cities Alliance, n.d., Ezeh et al., 2017). Therefore, the largest populations that make up slums are migrant workers, displaced persons, illegal and legal immigrants, unemployed persons and refugees (Riley et al., 2007). In the context of global population growth, people move from the countryside or other areas to cities to obtain better livelihood opportunities (Cities Alliance, n.d., Ezeh et al., 2017, Keivani and Werna, 2001, Okpala, 1992). These migrants do not necessarily have the skills to gain employment in the urban job market and can only be employed in the informal sector. Rural migrants are therefore deprived of higher incomes and formal sector benefits (Gundogan and Bicerli, 2009, Posada and Moreno-Monroy, 2017, Riley et al., 2007). Due to a lack of capital, and legal and policy recognition, these migrant’s living spaces are also outside the governmental development framework and planning. The result is the informal settlement or slum (Azunre et al., 2022, Okyere et al., 2017, Sheppard et al., 2020). Such settlements are often the first point of arrival and encounter with cities for rural migrants, providing asylum and further opportunities for formal status (Alvarado, 2022, Cities Alliance, n.d., Keivani and Werna, 2001, Marris, 1979, Misselhorn, 2008). This is also why Taubenböck et al. (2018) refer to such areas as "*Arrival Cities*".

It follows that water practices within slums differ significantly from other urban areas (Azunre et al., 2022), with the urban poor facing considerable difficulties in accessing water (Nyika and Dinka, 2023, Richmond et al., 2018, Sinharoy et al., 2019). Slum residents often rely on informal suppliers for water (Dakyaga et al., 2021). In Dhaka, Bangladesh, the urban poor rely on water vending and use water from dug wells, ponds, rivers, canals, and swamps (Akbar et al., 2007). In Luanda, Angola, water truck operators distribute water to informal home-based water retailers who sell to urban populations or water street vendors (Cain, 2018). Though operational methods vary, these unauthorized suppliers fill gaps in formal water supply within slums (Joshi et al., 2023). This gap, as summarized by Sinharoy et al. (2019) and Nyika and Dinka (2023), stems from high infrastructure construction costs, slum dwellers' low willingness and ability to pay, long distance of slums from key urban areas, land hazards, complex building environments, marginalisation of residents, insecure land tenure, and lack of legal and planning recognition. Specifically, Sultana (2020) points out the impact of informal citizenship on slum residents' access to water. Joshi et al. (2023) emphasize the influence of lack of tenure security as a reason that slum dwellers turn to informal water services. Insecure land ownership and building configurations in slums often lead to chaotic network structures that are difficult to maintain and manage (Ahlers et al., 2013, Boakye-Ansah et al., 2019). Therefore, issues of access to basic services such as water are not merely technical or infrastructural concerns; rather, they are closely intertwined with broader processes of urbanization and informality.

The significant role of informality in the economic, environmental, and social sustainability of cities in the Global South is increasingly being recognised (Azunre et al., 2021, Azunre et al., 2022). In particular, the informal sector meets the needs of those excluded from the postcolonial urbanisation process in a flexible matter, including slum dwellers and those considered to be from the lower classes. Soliman (2020), Azunre et al. (2022), Smit et al. (2017), and Azunre et al. (2021) point out that urban informal activities play a crucial role in sustaining residents' livelihoods and making significant economic contributions in cities across SSA and the Middle East. With appropriate management, informality can facilitate transitions toward sustainable development. Auerbach et al. (2018) and Caldeira (2017) emphasise the political impact of informal collectives on urban governance, enabling marginalised residents to assert their agency in production and decision-making processes. Traditional perspectives analyse informality and urbanisation in the Global South through the lens of northern urban experiences. For instance, Lawhon and Le Roux (2019) compared urban theories included in mainstream textbooks and highlight the tendency to treat northern urbanisation as paradigmatic, while considering urbanisation in southern cities as exceptions. The typical hegemonic narrative is that governance actions in southern urbanisation are seen as imitations of those in northern cities (de Castro Mazarro, 2023). Such a perspective fails to account for the informal practices that characterise urbanisation in the Global South (Roy, 2011, Sheppard et al., 2013).

Many opinions have arisen as to whether "formal" should be regarded as the norm and "informal" as abnormal in urban life. This distinction can be roughly seen as a difference in the understanding of northern and southern urbanisation (Brenner and Schmid, 2015, Sheppard et al., 2013). Unlike global urbanism, which takes economically dominant global cities as models for urban development, these alternative perspectives do not simply treat informality and inequality in Southern cities as problems to be overcome (Roy, 2011, Sheppard et al., 2013). Instead, they seek to identify similarities between urbanisation processes in both the Global South and North, and aim to conceptualise informality (Harris, 2020, Nijman, 2007). Such frameworks include subaltern urbanism (Moyo, 2023, Roy, 2011) and critical urban theory (Brenner, 2009). Subaltern urbanism interprets urbanisation from the perspective of the lower-class domain represented by subordinated slums. Drawing on critical urban theory, Banks et al. (2020) contend that informal practices under the dichotomous perspectives should instead be understood as adaptive strategies employed by differentiated social groups. This perspective helps to explain the emergence of new forms of “informality”, such as public-private partnership arrangements, which will be introduced in Section 2.2.3. This argument is also reflected in Moretto et al. (2018)’ study of water and sanitation service co-production in four case study cities. They concluded that informal services are not entirely opposed to formal services; instead, there is significant interaction between the two. Parida and Agrawal (2023) advocate for an approach that analyses southern cities based on their characteristics. This perspective aligns with the principles of comparative urbanism (Robinson, 2016), which emphasises the contributions of diverse urban experiences to urban theory.

It is important to note that acknowledging the role of the informal sector in developing countries does not imply ignoring the differences between formal and informal, but rather calls for a deeper understanding of their organisation (Caldeira, 2017, Sheppard et al., 2013). Nevertheless, these perspectives provide a way of understanding slums that has facilitated and justified the implementation of slum upgrading and water governance policies, which will be introduced in the following sections.

2.2.2 Policy towards slums: clearance versus upgrading

City authority policy towards slums, particularly slum housing and infrastructure, have varied. Before the 1970s, many governments showed little interest in improving slum conditions, instead opting for slum clearance programs to relocate residents to other settlements. Such initiatives included post-war reconstruction projects in UK cities (Jones, 2008, Kearns et al., 2019), the two-stage slum clearance program implemented by London County Council between 1889 to 1907, the Cross Act (1875), and Torrens Act (1866) (Stewart, 2005, Yelling, 1982). Other slum clearance projects can be found elsewhere, such as in the United States (Hill, 1952), Nigeria (Adama, 2020), and Zimbabwe (Muchadenyika, 2015). One of the most direct reasons for the launch of these projects was often financial gain. Since slums are formed because immigrants seek job opportunities in cities, the more convenient a slum's location is, the better it is for immigrants. Without regulation, they often occupy locations with easy transport access (Marris, 1979, Misselhorn, 2008). Thus, slum locations are attractive not only to immigrants but also to governments and developers—especially given that slums are often not protected by law (Caldeira, 2017, Cities Alliance, n.d., Jones, 2010, Li et al., 2023b, Viratkapan and Perera, 2006). Moreover, the objective of slum clearance is to reduce unfit housing, and such schemes assume that the only solution is to eradicate the slum and relocate the residents to other places (Andavarapu and Edelman, 2013, Yelling, 2000).

However, as indicated in section 2.2.1, slums should not be seen merely as obstacles to southern urbanisation. Thus, slum clearance policies have been widely criticized. A slum is a shelter for its inhabitants, not only in terms of the buildings and infrastructure present in the slum, but also in terms of informal organizations and activities. These organizations and activities result from continuous negotiation, gaming and cooperation among multiple actors (the state, the private sector and high- and low-income residents, etc.) in the urbanisation process (Lara-Hernandez et al., 2020). These informal activities and behaviours fill gaps that the formal order does not cover. Specifically, they provide services, employment opportunities, and other social environments upon which household and social networks depend (Misselhorn, 2008, Roy, 2011). Thus, moving slum residents out of slums is not only an emotional blow but also a blow to their livelihoods, making people who are already living precariously poorer; this is the '*destruction of communities*' theory that prevailed in the 1950s to the 1970s (Abebe and Hesselberg, 2015, Brown-Luthango et al., 2017, Kearns et al., 2019, Olu Sule, 1990). More broadly, such practices have been described as *urbicide*—the deliberate destruction of the material, cultural, and socio-economic foundations that support community life, often in the name of urban renewal (Di Virgilio, 2023).

In the 1960s and 1970s, scholars such as Turner commented on housing policy and put forward the theory of gradual improvement. As a representative of slum upgrading, Turner, based on his observations of Peru, proposed that if the government could improve the unsanitary environment in the slum, residents would gradually improve shanty dwellings spontaneously, so intervention by government bureaucracies should be limited (Werlin, 1999). His discourse influenced World Bank policy, which directly affected countries' attitudes towards slums and brought slum upgrading from theory to reality (Werlin, 1999). As the understanding of urban informality grew, the strategy of upgrading slums was promoted (Cities Alliance, 2021c). However, it should be noted that although the 1970s are commonly viewed as the starting point of the global slum upgrading trend, Harris (2020) reminds us that slum upgrading projects had already been implemented in both the Global South and North prior to the 1970s. These early efforts, though less documented and often overlooked in academic narratives, laid important groundwork for later upgrading strategies.

Although the forms of intervention vary significantly, slum upgrading projects generally recognise improvements to physical infrastructure (Adama, 2020, Brown-Luthango et al., 2017, Cities Alliance, 2021c). In particular, due to the critical importance of water, sanitation, and hygiene (WASH), their role in slum upgrading projects has been widely discussed, especially after the pandemic (Cities Alliance, 2021b, Olthuis et al., 2015). Consequently, many slum upgrading projects involve the provision of water infrastructure. The upgrading project in Bandung City, Indonesia, provides piped water to 121 urban villages with slums through communal boreholes (KOTAKU boreholes) (Urfanisa et al., 2022). In 2006, Nairobi, Kenya, launched the Urban Basic Services (UBS) project as a sub-project of the Kenya Slum Upgrading Programme (KENSUP) (GoK, 2004, Meredith and MacDonald, 2017). Additionally, Nairobi has other water and sanitation infrastructure upgrading projects, such as the Kibera Community Water and Sanitation Project (Cronin and Guthrie, 2011).

Another feature of slum upgrading is participation. City Alliance believes this is because of the occurrence of Turner's self-help housing concept (Cities Alliance, 2021c)， which de Castro Mazarro (2023) views as one of the outcomes of subaltern urbanism. As noted, informality is significant in Global South urbanisation, with entities in slums communicating directly with official authorities on issues of legalization, regulation, occupation, planning, and speculation (Caldeira, 2017). Such collaboration has the potential to improve public service provision (Chidambaram, 2020). Thus, community participation and active involvement in different phases of slum upgrading projects is considered important (Brown-Luthango et al., 2017, Svensson et al., 2003). Besides the mentioned UBS project in Nairobi (Meredith and MacDonald, 2017) and the slum upgrading project in Zimbabwe (Muchadenyika, 2015), slum upgrading in Bangladesh (Panday, 2020) and India (Chidambaram, 2020) has also highlighted the advantages of participatory slum upgrading and collective actions. Furthermore, participation in slum upgrading benefits the government via local knowledge – i.e. learning from informal activities (Brown-Luthango et al., 2017, Cities Alliance, 2021a, Nijman, 2008). The importance of community participation is widely acknowledged, especially in water and sanitation projects (Patel and Killemsetty, 2020).

In summary, better understanding of slums has led planners and practitioners to acknowledge the necessity for an approach that is more effective than clearance in addressing slum issues. New interventions are centred on infrastructure improvement, aiming to eliminate urban inequalities gradually by improving basic services such as water supply. Moreover, many advocates of slum upgrading believe in the active involvement of civil society, including the private and voluntary sectors (Nallathiga, 2012, Nijman, 2008, Otiso, 2003). Slum upgrading, particularly participatory slum upgrading, is a response to the traditional binary perspective on slums (Recio et al., 2017). Slum upgrading is an approach that views informality as a new mode of urban life and recognises informal efforts to improve basic services. The approach also implies that slums can play a positive role in sustainable urban development (Ahmed Saad et al., 2019, Azunre et al., 2021, Azunre et al., 2022).

2.2.3 Water service delivery policy: service co-production

Water utilities face distinct challenges in upgrading slums. The water sector naturally have a monopoly and public welfare remit. Their monopoly arises because of the high barriers to entry. Given the high construction costs of water infrastructure, water utilities are more exclusive than sectors such as electricity and communications (Cesar, 2019). Their public welfare remit arises from the necessity of ensuring universal service availability across the entire population. This obliges operators to take public responsibility and prevents them from solely pursuing self-interest, which is undoubtedly contrary to private operators' profit-seeking drivers (Marin, 2009a, Ruiz-Villaverde et al., 2018). Furthermore, water pipelines are underground assets, making them difficult to catalogue and monitor, thereby presenting unpredictable risks for private companies.

In order to alleviate public debt and curb high inflation in the 1970s, many countries significantly reduced subsidies available to the public sector (Ruiz-Villaverde et al., 2018). Given that tariff revenues were insufficient to cover these reduced subsidies, public utilities tended to depend more on government budget transfers than tariff revenues. Therefore, water utilities were effectively forced to focus more on infrastructure expansion than on maintenance and management in order to gain access to government funds. This led to inefficient and low-quality water services. During this period, investment from government decreased, while service price increases were not feasible because consumers were reluctant to pay more for deteriorating services (Marin, 2009a). It was during this period that the informal water sector grew rapidly, and the perception of informal provision changed (Post, 2022).

Unsurprisingly, an innovative approach called Public-Private Partnership (PPP) aroused interest. PPP is a service co-production mode, which has relatively broad definitions and forms. Different researchers and organizations have defined PPPs based on their disciplinary interests or responsibilities (Hodge and Greve, 2017, Wang et al., 2018). The Organisation for Economic Co-operation and Development (OECD) defines PPPs as long-term agreements that the government uses to delegate services to a private partner to improve the efficiency of the service and transfer some risk to the partner (OECD, n.d.). Several commonalities emerge from these definitions:

1. Cooperation. PPPs are created to address issues that require cooperation between the public and private sectors.
2. Shared risks and goals. The concept of sharing is included in almost all definitions. Public utilities are characterised by public welfare, high risks, and costs. PPP spreads the costs and risks among the public and private participants and ultimately devotes them to achieving the common goal of public welfare.
3. Long-term cooperation. The high risk and cost of public services require long-term cooperation for the private sector to recover costs (Hodge and Greve, 2017). Therefore, Wang et al. (2018) believe that long-term contracts are the only viable form of PPP.

The rule setter of the water PPP is the public sector, and the form through which rules are defined is as contracts or concessions. Private operators can partially or fully invest in constructing new assets or use skills and expertise to optimize and manage infrastructure (Ameyaw and Chan, 2016, Marin, 2009a). Thereby, the management of public services is delegated to private developers, allowing public and private sector entities to share infrastructure costs and provide affordable water services to consumers whilst balancing this against potential losses of revenue.

In addition to being benefit-oriented, PPPs can also take social responsibilities. Such non-conventional models often involve the participation of community or civil society organizations and can effectively provide water to urban slum populations (Kleemeier and Lockwood, 2012, PPP Authority, 2008). Public utilities including water and sanitation are more vulnerable to the spatial inequality between slums and formal settlements. Healthcare utilities are a typical example. Disadvantaged groups (in terms of region, income, social class, race, gender) tend to seek assistance from healthcare utilities more (Hart, 1971), but they are not attractive to private operators (Armah et al., 2018, Knox and Pacione, 1980). The phenomenon is known as the inverse care law: “The availability of good medical care tends to vary inversely with the need for it in the population served.” (Hart, 1971). Hart (1971) believes that market forces are the driving force of this law, and the stronger the privatization, the stronger the inverse effect. Likewise, slum residents are also not attractive to private water and sanitation operators. As noted in section 2.2.1, since slums tend to be located on the outskirts of cities and are perceived as having low willingness-to-pay for services, public sector bodies have little incentive to expand water and sanitation services in slums (Allen et al., 2006, Castro and Morel, 2008).

Therefore, PPP is considered as an effective approach for upgrading slum infrastructure, especially water infrastructure, as it can integrate local resources (Annamalai et al., 2016, Dos Santos et al., 2017, Lima et al., 2021, Marin, 2009b, Moretto et al., 2018, UN-Habitat, 2023b). Moretto et al. (2018) argue that small-scale customised solutions to water issues are available in PPP projects due to their closer connections with residents. Asumadu et al. (2023), on the other hand, highlighted PPP's superiority in financing. A widely studied example is the delegated management model (DMM) implemented in Kenya slums. In Kisumu, Kenya, Kisumu Water and Sanitation Company (KIWASCO) sells water through trunk pipelines to small-scale providers (SSPs), who then distribute water to consumers through pipelines, shared standpipes, and water kiosks (World Bank, 2009). In which SSPs can be community-based organisations or individuals (Castro and Morel, 2008). The DMM approach has expanded the coverage of water services while also providing better quality water to slum residents (Nzengya, 2018, Schwartz and Sanga, 2010). Additionally, community participation in DMM has not only alleviated the burden on the government but also reduced the cost of water access for slum residents (World Bank, 2009).

However, there is an ongoing debate about the performance of PPP projects in the water sector (Cesar, 2019, McDonald, 2018, Mvulirwenande et al., 2019, Ruiz-Villaverde et al., 2018). Andres et al. (2008) collected data on 45 private water and sanitation companies in Latin America. The results show that private companies have higher labour productivity, efficiency, and quality levels. However, it is impossible to tell whether this improvement is due to privatization or company management. Chenoweth and Bird (2018) reviewed 20 studies on water and sanitation services in the United States, England, Wales, and France. They concluded that there was no convincing evidence that private companies outperformed public sector entities. Kirkpatrick et al. (2006) interpreted data from water utilities in 13 countries. They suggest that the impact of privatization was positive across multiple service provision domains, but not statistically significant.

According to the World Bank's Private Participation in Infrastructure Project Database, the number and investment of PPP projects in 2023 have decreased compared to 2022 (WBG, 2024). Studies have identified more than 20 failure drivers facing water PPP projects, including corruption, illegal connection, political interference, infrastructure construction and maintenance failures, and unclear land ownership (Ameyaw and Chan, 2015, Zhang and Tariq, 2020). Furthermore, Lima et al. (2021) and Almeile et al. (2024) highlighted risk-sharing issues in PPP projects, as the large scale of investment and the long payback period put the great risk on the private sector and make them lose interest in PPPs. Additionally, poor contract arrangements can lead to the private sector having difficulties in fulfilling promises, resulting in early contract termination and a lack of subsequent bidding interest. Buenos Aires, Hamilton, and Dar es Salaam have all seen PPP projects cancelled because private operators are unwilling to comply with unprofitable and restrictive contracts (McDonald, 2018). Studies on southern cities also indicate that water PPPs often suffer from weak infrastructure, limited funding and technical resources (Adams et al., 2020). Additionally, the importance of legislative support and recognition is highlighted for the success of PPPs in India, Latin America, and the Caribbean (Jha, 2023, Munoz-Jofre et al., 2023, Tirumala et al., 2020). Despite the increasing research on water sector PPP projects in recent years (Lima et al., 2021), there is a notable lack of studies examining the risks encountered by PPPs in slums and their interactions with the slum physical environment (Henson et al., 2020), such as the constraints imposed by slums' spatial layout on DMM operation (Nzengya, 2015).

2.3 Impacts and drivers of water pipeline configuration on services in slums

2.3.1 Drivers of water pipeline configuration in cities

Whilst the challenges of slum upgrading and water delivery policy within slums are well known, their implications for configuration of the associated infrastructure, and how utility networks within slums differ from formal areas are rarely discussed.

Sorensen (2018) points out that, like railway and road networks, establishing water and sewer networks requires significant financial investment. Once the network is built, there is no way to move these capital investments. The presence of sunk costs leads to the following characteristics of pipeline networks: firstly, urban networks tend to be monopolised because the high costs preclude the possibility of free competition. The threshold is high, and the costs caused by competition are unaffordable; secondly, due to the high costs, there is a severe path dependency in the construction of the network. Once the location of the network has been determined initially, subsequent network expansion will be highly dependent on the location of the existing network. Similarly, it is difficult to change the management system overseeing infrastructure once established. These characteristics can be summarized by the Matthew effect: "the rich get richer and the poor get poorer" (Merton, 1968), which is also observed with road networks (Lan et al., 2022). From a topological perspective, this effect can be explained by the preferential attachment process or the Yule process, where a small number of nodes (hubs) assume the primary connectivity roles during network formation, and new nodes tend to connect to these highly connected hubs (Fornito et al., 2016b, Yule, 1925).

These characteristics can be traced back to the nineteenth century when large-scale urban pipeline networks emerged in industrial cities. The expansion of industrial cities was accompanied by high densities of population and human activity. In order to provide these cities with adequate services, centralised systems were needed to gather funds and land for urban planning (Gandy, 2004). At the same time, the emergence of pipeline networks can also be linked to new public health concepts and technologies. It can be argued that the rise of pipeline networks represents the Western concept of the ideal city as a unified, orderly city, with centralised and standardised networked infrastructure serving a wide range of residents (Coutard, 2008, Coutard and Rutherford, 2015). In other words, pipeline networks are not just infrastructure but also a set of values. Equitable allocation of services to users via a pipeline network embodies a value system in which the equal status of the users in the city is central, as discussed in the concept of global urbanism in Section 2.2.1.

This value of equality is also reflected in designing the pipeline network. Since pipeline networks are constructed to provide universal access to public services, growing population with water demand becomes the main driver of water pipeline expansion. Network expansion in cities can be broadly classified into area expansion and densification (Yang et al., 2017). The former is the process of building a network to cover new settlements, while the latter refers to the enhancement of water services in existing communities. In both expansions, the distribution/density of population growth (demand points) determines the shape of pipeline networks (Farmani and Butler, 2014, Yang et al., 2017). These expansions align within the broader processes of urban evolution (Gudmundsson and Mohajeri, 2013, Mohajeri and Gudmundsson, 2014, Strano et al., 2012).

Consequently, pipeline networks exhibit similarities to the urban road networks, the latter of which fundamentally determine the morphology of cities (Scheer, 2015). Beyond their geometric overlap (Mair et al., 2017), graph studies also found shared patterns within urban networks, one of which is the power law distribution of centrality metrics (Akbarzadeh et al., 2018, Giudicianni et al., 2018, Johnson et al., 2019, Kirkley et al., 2018, Klinkhamer et al., 2017, Lämmer et al., 2006, Yang et al., 2017, Yu et al., 2024, Zischg et al., 2019). The power-law distribution is characterized by heavy tails, meaning that nodes with high centrality (i.e., the importance or influence of a node) are more likely to occur in scale-free networks than in Gaussian-distributed networks (random networks). Based on the definitions of centrality metrics, the rule reflects the pattern shared by urban sprawl and infrastructure networks, where the network's primary connections (the city's framework) form circuits and then expand outwards (Akbarzadeh et al., 2018). The evolution of both pipeline and road networks both follow this development dynamic (Dovey et al., 2020, Mohajeri et al., 2015).

Although many have questioned the universality of power law distributions in recent years (Broido and Clauset, 2019, Casali and Heinimann, 2019, Reza et al., 2024, Wéber et al., 2020), there is evidence that a significant proportion of real networks are subject to this unifying law or alternative distributions, such as lognormal (Casali and Heinimann, 2019) and modified Lomax (Akbarzadeh et al., 2018, Artico et al., 2020, Chattopadhyay et al., 2021, Kirkley et al., 2018). Graph metrics have revealed statistical similarities between road networks across urban contexts (Akbarzadeh et al., 2018, Casali and Heinimann, 2019, Kirkley et al., 2018, Lan et al., 2022). Liu et al. (2016) observed that building density is positively correlated with the centrality of streets. Giudicianni et al. (2018) and Krueger et al. (2017) identified regular variations in structural indicators of pipeline networks over time and urban scale. As both street and pipeline networks are constrained by the spatial boundaries established during city expansion, they tend to follow similar evolution trajectories (Abdel-Mottaleb and Zhang, 2020).

The research described above has taken place mainly in developed countries and has focused on the formation patterns of general urban networks. In Africa, however, the environment affecting the layout of pipeline networks is more complex. When this centrally managed pipeline system was transplanted to Africa from colonial times, the lack of resources limited the network's layout and, thus, its performance. As Gandy (2004) states, contrary to what one would expect, not all cities will follow the western network pattern. The first issue many African cities face is the profound influence of colonial policies (Andersson, 2017, Harris, 2021, Letema et al., 2014). During the colonial era, African colonies were often divided into two zones: *la ville des indigenes* (the indigenous zone) and *la ville des europèenes* (the European zone) (Bigon and Njoh, 2015). This stratification was based not only on ethnicity but also on native locations, culture, religion, occupation, and income. In Tema, Ghana, for example, colonial housing allocation policies were designed according to occupation and income (Kaye-Essien, 2020). The result of stratification was the explicit spatial and functional divisions within African cities, whereby the native areas severely lacked urban planning and infrastructure services compared to the European areas (Tetteh et al., 2022). This unfair dualistic structure has been maintained during subsequent urban development and makes building equally distributed pipelines difficult in African cities from the outset (Bigon and Njoh, 2015).

As mentioned previously (section 2.2), the second factor influencing pipeline networks in African cities is rapid urbanisation and population growth, accompanied by greater water service demand. Individual local authorities or utilities in African cities lack sufficient funds and the construction and management skills necessary to build and manage a pipeline network that covers all demand equally. Instead, the dualistic structure can still be found as a legacy in water pipeline networks even after African countries gained their independence (Bigon and Njoh, 2015). In the case of Kisumu, for example, although population density is the criteria for prioritising pipeline installation, it is the formerly European or Asian middle and high-income areas with lower population density rather than slums and peri-urban areas that are included in planning (Letema et al., 2014). Similar problems are seen in Lima, Peru, and Dar es Salaam, Tanzania (Ioris, 2012, Smiley, 2020). Further, Ioris (2012) concludes that in a context where the authorities lack the capacity to manage the whole city, water supply shortages result from colonial and post-colonial social, economic, and spatial inequalities. Jaglin (2012) emphasises that social inequalities have caused different governance arrangements for particular population groups and areas. This contradicts the values on which infrastructure networks are based - equality and universal access (Coutard, 2008). As Dupuy (2011) states, "Even though the network exists, all the necessary elements for it to function are not yet in place."

Predictably, if one compares the morphology of pipeline networks within slums with that of formal areas, they will be very different since:

1. The unplanned buildings and road networks in slums limit the space available for pipeline laying, thereby making the pipelines have inefficient topology.
2. Insecure land ownership, low-income levels, and governance failures further weaken pipeline planning and maintenance in slums.
3. Due to colonial history, the infrastructure baseline in slums differs significantly from that in formal areas, increasing the cost of further pipeline expansion.

Several empirical studies have confirmed this argument (Lagerberg, 2016, Mapunda et al., 2018, Shushu et al., 2021). Due to their populations' inability to access and afford formal water services, slums suffer from illegal connection problems and thus have unplanned piped network structures (Boakye-Ansah et al., 2019). For example, Mutikanga et al. (2009) reported the formation of a "spaghetti" pipeline network, referring to a structurally disorganised network (see Figure 2.3), in Kampala, Uganda. Another reason for slums' vulnerable networks is that urban expansion is faster than expectations. Therefore, these networks are built without utility plans (Shushu et al., 2021). High building density, the complex road network, and insecure tenure also inhibit building of pipeline networks in slums (Alba and Bruns, 2022, Wagle, 2022). It is worth noting that there are no studies yet on the differences in the structure of pipeline networks within slums versus formal areas. In addition, although some interventions (such as DMM) have been implemented in slums to improve water access (Nzengya, 2018, Ocholla et al., 2022), their impact on network structure remains unclear (Nzengya, 2015).

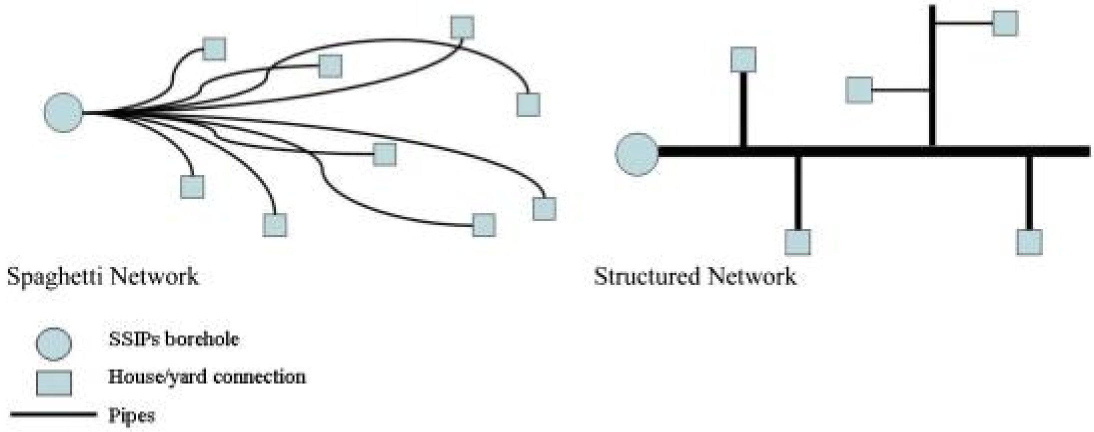


Figure 2.3 Spaghetti networks versus structured networks. The network endpoints are Small Scale Infrastructure Providers (SSIPs) and households (Ahlers et al., 2013).

2.3.2 Impacts of water pipeline configuration in cities

As discussed in the previous section, the urban context shapes the configuration of pipeline networks, creating a contrast between informal and formal areas. A related question is whether pipeline performance varies as a result of their varying configurations in different contexts.

Here, the review first discusses reliability as an example, which refers to the ability of the pipeline network to remain operational during malfunctions (Gheisi et al., 2016). Pipeline networks and other infrastructure networks are built to transport substances and energy from one point to another. Given equity is an underpinning value of many networked cities (see Section 2.3.1), it is crucial that users have equal access to substances and energy from the network. This is also consistent with the universal access targets for SDG 6. Therefore, the reliability of a network can be measured by the number of off-grid customers when failure happens. In other words, it is the risk that when one pipeline fails, other pipelines are disconnected from the network, which is affected by how the pipelines are organised (Agathokleous et al., 2017, Bentes et al., 2011, Wang et al., 2019). According to Punmia et al. (1995), the way pipelines are organised can be classified into four types:

1. Grid-iron system or Reticulation system
2. Circular system or Ring system
3. Radial system
4. Dead-end system or Tree system

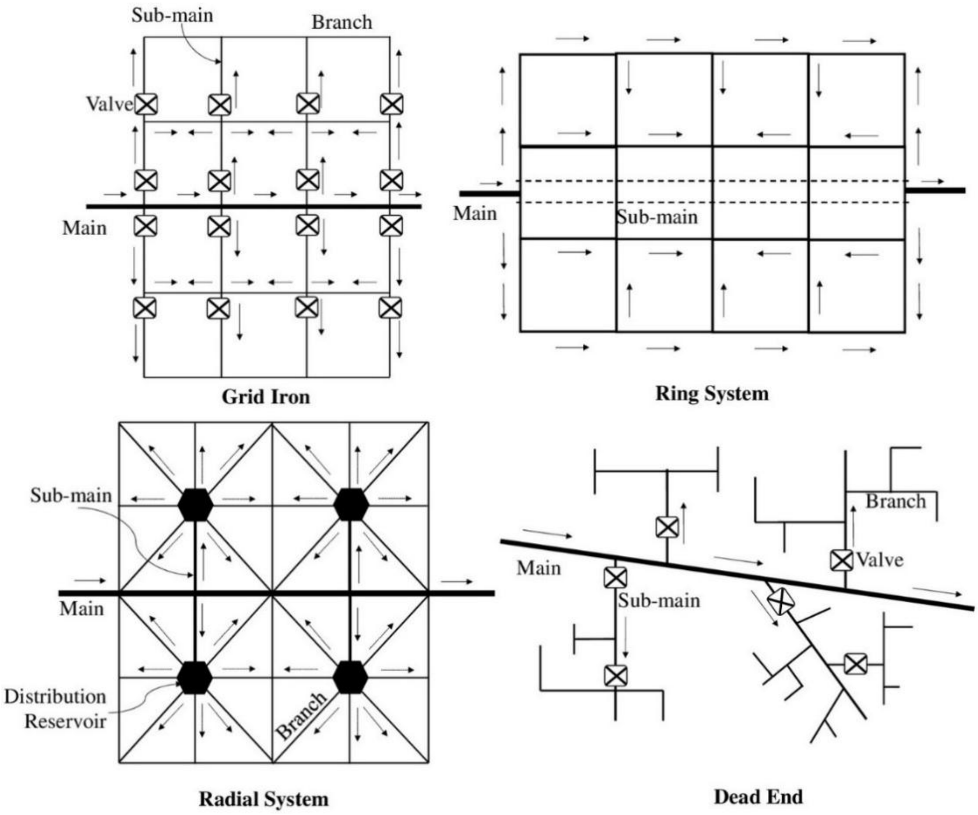


Figure 2.4 Water distribution system layouts. A. Grid Iron; B. Ring System; C. Radial System; D. Dead End (Mazumder et al., 2018).

These layouts appear in different forms of settlements and have different advantages and disadvantages. In the case of a dead-end or tree system (Figure 2.4 (D)), which is common in rural areas (Sarbu and Ostafe, 2016), a main pipeline runs through the service area with sub-mains connected on either side. These sub-mains are further divided into branch lines, where there are no cross-connections between sub-mains and branch lines. The whole system forms the shape of a tree, and the trunk is the main pipeline. In this network, breakage of any pipelines will result in some customers not receiving service, and the extent of the impact depends on the class of the pipeline that fails. The spaghetti network mentioned in section 2.3.1 fits into this category. However, this does not imply that networks within slums are necessarily dead-end systems, as slums exhibit various spatial structures (Flores Fernandez, 2011). On the contrary, all branches are connected in the ring systems (see Figure 2.4 (B)) that can be found in well-planned cities. Very few customers in this system will be affected when one pipeline breaks because there are always several pipes connecting a point. Lorenz et al. (2021) suggest that the resilience of pipeline networks varies with urban form, as urban form greatly affects the structure of pipeline networks. Moreover, Adraoui et al. (2024) suggest that it is also possible to estimate the robustness of a network by testing the consequences of removing its constituent pipelines using graph metrics. Since the distribution of centrality can be used to describe the network's structure (see section 2.3.1), some studies argue that the power law distribution of a pipeline network implies the existence of highly concentrated centres. Such a scale-free network is highly resilient to random faults compared to a random network, but vulnerable to targeted attacks. Thus, the degree to which the distribution is subjected to a power law can be a useful indicator of network resilience (Wéber et al., 2020, Yu et al., 2024, Zischg et al., 2019).

In addition to reliability, water pipeline networks have unique properties that are affected by structure. Many studies use graph metrics to measure the impact of a network's topology on the transmission efficiency of water flow or, put another way, the importance of nodes/edges and the robustness of the network. Yu et al. (2024) demonstrate that the topology of the water network can be used to estimate hydraulic head, flow rate, water age/residency times, and thereby water quality in the network. Zhao et al. (2020) found that energy consumption for water supply (e.g., energy required for pumping stations) varies with urban morphology, with radial-uniform cities consuming the least energy for water supply among radial, grid and satellite cities. Torres et al. (2017) pointed out that changes in network topology lead to subsequent changes in water quality in the system. He examined 11 graph metrics regarding maximum hourly unit head loss, minimum hourly system pressure, average system water age, and average concentration of an unknown chemical. The result shows that 10 of the 11 metrics describing network structure and 3 of the 4 performance metrics had strong positive or negative correlations with these operational water parameters.

Based on the evidence from section 2.3.1, relationships between piped network performance and urban form can be further inferred. If the performance of a pipeline network is influenced by its structure, then a pipeline network’s weak structure could exacerbate water service delivery challenges in slums. Taking Water Safety Plans (WSPs) as an example, a WSP is a management tool designed to ensure the delivery of safe drinking water (Roeger and Tavares, 2018). It enables managers to effectively identify and control risks within the pipeline network. The development of a WSP relies on quantitative analysis and supporting programmes to oversee the water system (Godfrey and Howard, 2005b). However, Godfrey and Howard (2005a) note that unplanned water systems with limited data availability make it difficult to locate the main pipelines. Consequently, the development of WSPs requires significant time and expert experience to conduct field surveys and desk-based system analysis, which delays preventative action against issues affecting water supply and quality. Additionally, for water pressure management, the water pressure, pipeline lifespan, and the water lost in leaks are closely linked. Thus, efficient water pressure management contributes to the system's short and long-term operation. However, in a ‘spaghetti’ network, managing water pressure in different zones is difficult due to the lack of information about infrastructure such as valves. Hence, frequent pressure transients and surges in the system will trigger bursts and leaks (Mutikanga et al., 2009). Similarly, Lagerberg (2016) notes that extensions of pipelines in informal settlements are often informal and, therefore, not officially documented. Especially in places like slums, where governance is fragmented or limited, few people know the pipelines' location and condition except those who laid them. Therefore, in the event of a pipeline breakage, few people realise when breakages occur and are able to remediate, repair or replace the pipeline.

Since maintaining a chaotic pipeline network is more complicated and costly than a well-structured network, bursts and leaks also lead to more severe consequences in a pipeline network with poorly structured topology. A complex pipeline network requires more computation to analyse the condition of pipelines (Ghosn et al., 2016). In addition, their size and complexity, as well as their accessibility, make repairing the system more difficult (Kaminsky and Kumpel, 2018). Particularly, pipelines in slums may be exposed rather than buried and made of plastic, allowing residents to cut the pipes to fetch water (Mapunda et al., 2018). An example comes from the slums of Malawi, where the maintenance cost per unit length of the pipeline is much higher than in other areas. The reason for this is the poorly laid network and the frequent vandalism of pipelines in slums (Banda and Mwale, 2018). In the case of Mwanza, Tanzania, the researchers observed that due to such bursts and leaks, the amount of non-revenue water in slum areas was much higher than the average for the city, with 87 per cent of the actual water loss being related to the network topology and operating conditions (Shushu et al., 2021).

This chapter underscores the close interconnection between SDG 6 and SDG 11, as revealed through topological research. The construction of water pipeline networks is fundamentally designed to provide equitable services to residents, aligning closely with the objectives of SDG 6 on safely managed water services. The realization of this goal is directly influenced by the structure of pipeline networks. Topological research finds that the ability of pipeline networks to reliably deliver water without being affected by failures depends on their structural type. Moreover, the network's structure directly impacts its hydraulic efficiency and ultimately parameters such as water residency times and pressure linked to water safety. Structured water pipeline networks are easier to manage, and thus have greater resilience against contamination and degradation. As a result, the structure of pipeline networks can provide insights into both aspects of the safely managed drinking water indicator of SDG 6.1: availability when needed and freedom from contamination.

On the other hand, similar to the informality of slums, spatial disparities in pipeline configurations across SSA also arise from urbanisation and are closely linked to urban structure. SSA cities face constraints such as limited financial resources and governance capacities, which curtail the extension of formal water services. Simultaneously, urban water sectors are under pressure to meet water demand due to population growth. A similar supply-demand contradiction is also the cause of slum proliferation, creating substantial overlap between water supply issues and slum challenges. Given that the evolution of pipeline networks mirrors urban development, significant disparities between slums and formal settlements in pipeline configurations should become evident through comparisons of urban and pipeline morphology. Thus, the subsequent empirical evidence in this thesis on how slum and slum upgrading projects affect pipeline network structure can inform efforts to achieve both SDG 6 and SDG 11.

2.4 Graph theory methods for measuring network structure

As previously highlighted, an approach often used in studies of the performance (e.g., robustness and efficiency) of urban pipeline networks is graph theory. Unlike hydraulic models, which measure the physical properties of pipeline networks (D' Ambrosio et al., 2015), graph theory measures the structure of the pipeline network. Its lower computational complexity makes it suitable for use in settings where data are sparse, which is particularly valuable given that the physical layouts (e.g., pipe layout, pipe diameter) and operational characteristics (e.g., pump operation) of pipeline networks are often not publicly available (Ahmad et al., 2022). Applications of graph theory include a range of metrics and algorithms that reveal infrastructure network characteristics from a different perspective, complementing traditional hydraulic models (Torres et al., 2017).

2.4.1 Graph approaches to water pipeline networks

A graph is a mathematical abstraction that can represent any set of elements related to each other in some way (Clark and Holton, 1991, Wilson, 2010). Loosely speaking, any system that connects individual units can be called a network and can be represented by a graph. A well-known graph example is Zachary's karate club. In this study, club members are treated as units, and the information flows between members are represented as graph edges (Zachary, 1977). Therefore, representing infrastructure such as pipeline networks as graphs is a natural idea.

There are various methods to represent pipeline networks as graphs. A pipeline network consists of components located in Euclidean space, with water flows forming connections that can also be mapped within this space. Such networks are known as spatial networks (Barthélemy, 2011, Tsiotas and Polyzos, 2018). Representing spatial networks as graphs inevitably involves discarding redundant elements while preserving their morphology and connectivity (Šuba et al., 2016). In other words, graph theory analysis requires the merging and simplification of spatial data to achieve a balance between mathematical representation and the real world, addressing various objectives such as aesthetics, readability, and computational efficiency. Since there is no universally accepted definition (Pueyo et al., 2019), the review uses the term 'graph generalization' here. Disagreements about generalization methods primarily revolve around the choice of which elements to discard and how to organise the components, which often depends on the research objectives (Pueyo et al., 2019). From a functional perspective, there are two generalisation methods for representing spatial networks as graphs. One approach is to depict components in the network that consume, generate, or regulate a resource or service as nodes (e.g., cities, ports), and the exchange of resources and services between nodes as edges (Dunn et al., 2013). The other method represents the intersections and endpoints in the network as nodes. Both methods can be seen in the study by Prieto-Curiel et al. (2022). Additionally, various methods have been proposed to identify and aggregate elements within spatial networks. Space syntax is a pioneering approach in this field. It represents linear elements in urban spaces as axial lines based on visibility (Hillier et al., 1993). Subsequently, methods have emerged for identifying elements based on features such as names and angular relationships between elements. These methods derive from specific interpretations of networks and have inherent limitations (Marshall et al., 2018). For instance, Marshall (2016) notes that while all nodes in air transport facilitate point-to-point services, in road networks, services do not terminate at some nodes. Consequently, the nodes have different functions, but the aforementioned representations fail to capture this distinction. Nevertheless, these modelling approaches all capture network continuity and hierarchical structure (Marshall et al., 2018, Negadi et al., 2023), and thus have also been applied to pipeline network modelling (Krueger et al., 2017, Zischg et al., 2019). However, the impact of using different generalisation methods in pre-processing pipeline network for analysis has yet to be fully recognised.

In this section, we discuss modelling pipeline networks as simple graphs. Generally, a graph without self-loops (edges connecting a node back to itself) or multiple connections between two nodes (parallel edges or multi-edges) is called a simple graph. Following the methods used in road network modelling, pipeline endpoints and intersections are normally represented as graph nodes, while pipes are represented as graph edges (see Figure 2.5).

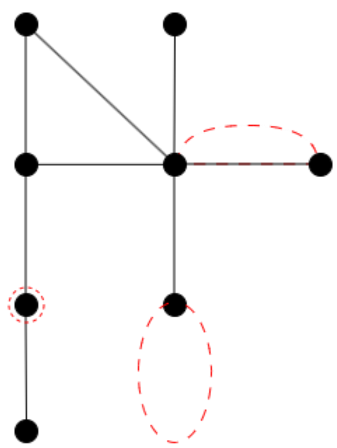


Figure 2.5 Example Graph of a Water Distribution Network. Dashed edges represent self-loops and parallel edges, while nodes surrounded by circles are pseudo-nodes.

Thus, a pipeline network can be represented by the adjacency matrix A. When the network has N nodes and E edges, A_{ij} is an *n* * *n* matrix and can be defined as (Barthélemy, 2011):

$$A_{ij} = \begin{cases} 1, & if\ i\ and\ \ j\ are\ \ connected \\ 0, & otherwise \end{cases}$$

For undirected and unweighted graphs, $A_{ij}$is a symmetric matrix, that is, A_{ij}=A_{ji} , which is not the case for directed graphs. For directed graphs, when i is connected to *j*, A_{ij} =1, but *j* is not necessarily connected to *i*, and A_{ji} may be 0. When the distance between nodes (whether physical distance or cost) also needs to be considered, we need to use a weighted graph to represent the network. For a weighted graph, when *i* and *j* are connected, A_{ij} =w_{ij} , where w_{ij}represents the distance of the edge connecting *i* and *j*. It can be seen that the adjacency matrix describes how the essential components of the network, the nodes, are connected.

According to this most basic definition of graph theory, we can already obtain some network metrics. Some commonly used metrics in water network studies are represented in Table 2.4.

Table 2.4 Basic Structural Network Measurements for Water Networks (*m* and *n* denote the number of edges and nodes in the graph, respectively, and *s*, *t*, *i*, *j* represent nodes in the node set *V*)

Graph metric		Formula	Description	Application in infrastructure network management
Centrality metrics	Degree centrality	$k_i = \sum_j A_{ij}$	The degree of a node is the number of its neighbours.	Centrality metrics indicate the importance of the node. There are a large number of them, including maximal clique centrality (MCC), maximum neighbourhood component (MNC), the density of maximum neighbourhood component (DMNC), betweenness centrality, bottleneck, eccentricity, stress, and closeness centrality (Barthélemy, 2011, Mata, 2020). Of these, degree, betweenness and closeness centrality are commonly used by recent infrastructure network studies (Daniel et al., 2021, Henry et al., 2019, Liu et al., 2016, Morzy et al., 2016). They assess nodes in terms of connectivity, flow loading, and efficiency, respectively. Their statistics
	Betweenness centrality	$B(i) = \sum_{s \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$	A node's betweenness centrality $B(i)$ can be defined based on the number of shortest paths through the node. σ_{st} is the number of shortest paths going from s to t and$\sigma_{st}(i)$ is the number of shortest paths going from s to t through node i. $B(i)$ can be interpreted as the ability of node *i* to facilitate the flow of material or information in the network.	
	Closeness centrality	$C(i) = \frac{1}{\sum_{i \neq j} d_{ij}}$	Closeness centrality measures the distance from a node to all other nodes in the network. $C(i)$ is the inverse of the sum of distances to all reachable nodes; the higher $C(i)$ is, the more efficiently the node can communicate with other nodes in the network.	

			also provide insight into the configuration of the network.
Link density	$q = \frac{2m}{n(n-1)}$	Density is the ratio of the actual number of edges in the network to this maximum possible and describes how connected a network is (O' Sullivan, 2014) .	Graph density represents how many edges can still be added to the network. Therefore, link density is the design efficiency of the network.
Critical breakdown ratio	$f_c = \frac{1}{\frac{k^2}{k} - 1}$	A network will fail when removing nodes whose degree exceeds the threshold f_c . The value thus depends on the average node degree k (Yazdani et al., 2013) .	This value is related to the stability of the pipeline network. The higher the value, the more tolerant the pipe network is to breakages represented via disconnected nodes.
Average shortest path length	$\overline{\ell_{ij}} = \frac{1}{n(n-1)} \sum_{i \neq j} \ell_{ij}$	The value of the average distance along the shortest paths ℓij connecting nodes i and j, compared to all possible pairs of nodes in the network (Porse and Lund, 2016, Yazdani et al., 2013).	This value is similar to link density. By traversing all node connections in the piped network, it represents the connection efficiency of the piped network.
Algebraic connectivity	λ_2	The second smallest eigenvalue of the normalized Laplacian matrix.	Similar to the critical breakdown ratio, algebraic connectivity represents the robustness of the pipeline network (Yazdani et al., 2013).

Meshedness	$\alpha = \frac{m-n+1}{2n-5}$	Also called the Alpha Index. Measures the ratio of the actual versus possible number of independent loops in a planar graph. It ranges between 0 for tree-like and 1 for grid-like networks (Porse and Lund, 2016).	A larger α corresponds to a more connected network (Hwang and Lansey, 2017) . This index can be used to describe the number of enclosed faces in infrastructure networks, thereby reflecting the structural characteristics of urban blocks(Usui and Asami, 2011).
Graph diameter	$\mathrm{Max}(d(N_i,\ N_j))$	The maximum geodesic distance between any two nodes. It captures the maximum eccentricity of nodes in the network and provides a basic measure of topological and geographical spread of the network (Zeng et al., 2017).	A rough estimate of the size and complexity of the piped network can be obtained from graph diameter.
Network efficiency	$\bar{Q} = \frac{1}{n(n-1)}\sum_{i\neq i} Q_{ij}$	Q is calculated as the ratio between physical distance and topological distance. Ranges between 0 for least-efficient and 100% for most-efficient networks and may be used as a proxy for average water travel time.	This is an indicator based on the actual operation of the pipe network. The higher the value, the less efficient the water flows in the network.
Central-point dominance	$C_B' = \frac{1}{\lvert V\rvert - 1}\sum_v C_B(v^*) - C_B(v)$	Central-point dominance is a parameter based on the evolution of betweenness centrality(Freeman, 1977). C_B' ranges from 0 to 1, and a higher value means that the nodes in the network are more closely	In a water network, a star graph can be effective in improving efficiency, meaning that there is a node in the centre of the network that plays an important role in

		distributed around a central point. 0 means that the distribution pattern of the points in the network has no distinctive features, while 1 means that the network is a wheel or star graph, which means that there is a significant central point in the network.	transmission, but this can also reduce the stability of the network, so in practice a grid structure is preferred for a water network (Yazdani and Jeffrey, 2011).
Clustering coefficient	$C_i^w = \frac{1}{s_i(k_i-1)}\sum_{j,h}\frac{w_{ij}+w_{ih}}{2}a_{ij}a_{ih}a_{jh}$	Also known as the Transitivity. Where $s_i$is the strength of vertex i, $a_{ij}$are elements of the adjacency matrix, $k_i$is the vertex degree, $w_{ij}$are the weights (Barrat et al., 2004). This metric detects the presence of triangular loops in the network. A value of 0 for C_i^w represents the absence of triangles in a network, whereas increasing triangle density results in a higher cluster coefficient. Therefore, grid networks have a smaller clustering coefficient.	Higher values indicate a more connected network and better performance in terms of network efficiency and redundancy. However, most loops in urban networks are not triangular but square. Therefore, another similar metric, meshedness coefficient, is always considered alongside clustering coefficient in water network studies (Yazdani and Jeffrey, 2011).

Spectral gap	$\Delta\lambda$	Spectral gap is the difference between the first and second largest eigenvalues of the adjacency matrix.	The magnitude of this value is related to a property known as "expansibility". Intuitively, it represents the connectivity and robustness between any set of points in the network. Low values of spectral gap indicate a lack of expansibility and are more prone to failures when the network is under attack (Yazdani and Jeffrey, 2010) .

2.4.2 Community detection for identifying network sub-regions

In addition to examining the attributes of urban pipeline networks, the relationship between pipeline network performance and spatial patterns implies that urban heterogeneity, or the distribution of slums, affects water delivery within sub-regions. Further, given the interactions between SDG 11 with SDG 6, a tool such as community detection can be used to identify the morphology of the water pipeline network and thus evaluate whether the piped structure for water delivery within slums and areas under specific service delivery management arrangements differ to those elsewhere.

Community detection is a technique for identifying communities—nodes in a network that are tightly connected or share similar characteristics (Fortunato and Newman, 2022). This concept builds on the ideal network. In mathematics, we can build and elaborate on ideal models to test similar model properties in more complex cases. This is also the case with the network model. A popular approach is to define random graph classes. One null model that is important for any network model is the Erdös-Rényi (E-R) model (Erdős and Rényi, 1960). The E-R graph is built from a set of random nodes, where the probability p of connecting two nodes is the same.

While the E-R model is idealised, real-world networks tend to be highly heterogeneous. Let us first return to the study of Zachary's karate club (section 2.4.1), which is not only a landmark study of social networks, but also inspired the study of community structure in the network. The karate club has become the test subject of many community detection algorithms (Chintalapudi and Prasad, 2015, Despalatović et al., 2014, Girvan and Newman, 2002). During Zachary (1977) 's study, the club was divided into two parts due to conflict between members. Members in each part have stronger social ties to each other and fewer ties between different parts. From this example, it can be recognised that in a social network where each node represents a person, the connection between nodes, or the social relationship, is affected by people's preferences. In addition to Zachary's karate club, a study on a collaboration network of scientists at the Santa Fe Institute, an interdisciplinary research centre in Santa Fe, New Mexico, showed that scientists tend to work with colleagues with similar research themes or methods (Girvan and Newman, 2002).

The above examples demonstrate a property that occurs in real-world networks: the existence of a set of entities that are closer to each other than other entities in the dataset, called a community. This means that a network influenced by or consisting of people is not a random graph like the E-R graph. The community structure in networks has been studied extensively in sociology (Kao and Porter, 2018) and can even be extended to biological research (Girvan and Newman, 2002). In addition, there are also community studies in the fields of land use (Comber et al., 2012), transportation networks (Hong and Yao, 2019), and water pipelines (Brentan et al., 2017). Studies have revealed the characteristics of communities and, further, how they are influenced by entities within networks (Javed et al., 2018). For example, the process of disease transmission can be understood by identifying different population communities (Kitchovitch and Liò, 2011), recommendation algorithms can be improved for business websites (Krishna Reddy et al., 2002), and the division of brain regions can be studied (Zhang, 2017). In the fields of communication networks, economics, and biological networks, community detection has also produced many outputs (Fortunato, 2010, Javed et al., 2018, Mittal and Bhatia, 2021).

However, it is not easy to review community detection algorithms. Due to the complexity of real-world networks, many algorithms are developed for specific communities and network types (Yang et al., 2016), which means that community detection algorithms are best classified relative to a specific network and community application area. For example, Mittal and Bhatia (2021) classify community detection algorithms into four categories according to technology: modularity algorithms, information theoretic algorithms, network algorithms, and hierarchical algorithms; while in Bedi and Sharma (2016)'s review, the algorithms for community detection are categorized into approaches based on graph partitioning, clustering, genetic algorithms, label propagation-based, semantics-based, methods for overlapping community detection, and community detection for dynamic networks. Another algorithm classification differentiates the following categories: traditional algorithms of community detection, algorithms of overlapping community detection, and algorithms of local community detection (Wang et al., 2015). While it is difficult to categorize algorithms, we can still filter out some algorithms since the type of network and community largely limits their application.

Networks can be divided into static and dynamic. Community detection algorithms in static networks are relatively simple and mostly centred on optimizing the objective function (Javed et al., 2018). However, networks in real life may be dynamic, which is reflected in the changes in network structure and composition. For example, Twitter users may be interested in different topics at different times and thus join different communities. In addition, communities can be overlapping, meaning that members of a community can simultaneously belong to another community (Mittal and Bhatia, 2021). Given that urban area divisions typically do not have overlapping or ambiguous regions, and that the analysis of the evolution of water networks is not within the scope of this review, the review focuses only on algorithms for non-overlapping communities and static networks. The following section is an introduction to three common types of algorithms.

2.4.2.1 The Cut-based Perspective

Since communities have stronger internal connections than external connections, some algorithms have found a way to delineate optimal boundaries to cut the graph so that the internal connections of the cut result are stronger than those crossing boundaries. In the traditional partitional clustering algorithm, given the number of target clusters, the performance of the given cost function is continuously optimized during the cutting process, such as maximizing or minimizing a loss function based on the distance between clusters. Functions such as Minimum k-clustering, k-clustering sum, k-centre, and k-median are all this type of community detection algorithm (Wang et al., 2015). They are all functions assigning points to k given clusters by optimising the distance between the k centroids and the other points within the cluster. The result is Voronoi cells. However, since it is difficult to know the number of clusters in advance when conducting community detection, hierarchical clustering methods are often proposed as an alternative to such methods. These algorithms consider the network as a binary tree with different levels. There are two branches of hierarchical clustering methods. The starting point of agglomerative algorithms, the first branch, is nodes. At any step of an agglomerative algorithm, whether two nodes are connected depends on the similarity score of the cluster. This method is an iterative process of merging communities from the bottom up (Despalatović et al., 2014). In contrast, divisive algorithms are global in perspective, consider all nodes as one community, then iteratively split clusters top-down by removing edges connecting vertices with low similarity. Both algorithms are iterative, as the weights of the edges change after each operation. Among them, divisive algorithms belong to the cut-based perspective. A typical example is the Girvan-Newman method, where edge weight is defined as the number of shortest paths passing through a given edge, a value known as edge betweenness (Girvan and Newman, 2002).

Graph Partition methods, on the other hand, aims to partition a graph into multiple predetermined-sized communities that satisfy some objective function by removing edges. A typical method is the Kernighan-Lin algorithm proposed to deal with circuit problems (Kernighan and Lin, 1970). This is a greedy optimization algorithm whose basic idea is to maximize the profit function by exchanging nodes between different groups. This profit function can be defined as the difference between the number of edges inside the module and the number of edges lying between them. Nevertheless, this algorithm requires the size of the community to be known. If the number of communities is known, then another algorithm in graph partition, spectral clustering, can be applied.

Similar to partitional clustering, if the sum of the weights of the edges between different clusters is the cost (rather than the distance between clusters), the basic principle of spectral clustering can be obtained (von Luxburg, 2007).

For two node sets $A, B \subset \mathrm{V}, \mathrm{A} \cap \mathrm{B} = \emptyset$, the cut weight between *A* and *B* is:

$$W(A,B) = \sum_{i \in A, j \in B} w_{ij}$$

When cutting the graph, for a set *k* of subgraph points A_1, A_2,.. A_k, the cut is:

$$cut(A_1, A_2, \dots A_k) = \frac{1}{2}\sum_{i=1}^{k} W(A_i, \overline{A_\iota})$$

Where $\overline{A_\iota}$ is the complement of A_i.

If the calculation process stops here, the result is likely to be inaccurate: isolated points in the network will be classified as clusters. Therefore, spectral clustering introduces the Laplace matrix to ensure the size of the clusters is sufficiently large. Take the RatioCut algorithm as an example (Wei and Cheng, 1989):

$$RatioCut(A_1, A_2, \dots A_k) = \frac{1}{2}\sum_{i=1}^{k} \frac{W(A_i, \overline{A_\iota})}{|A_i|}$$

This equation takes into account the size of the clusters.

At the same time, it introduces the indicator vector $h_j \in \{h_1, h_{2,\dots} h_k\}$:

$$h_{ij} = \begin{cases} 0 & v_i \notin \boldsymbol{A_j} \\ \dfrac{1}{\sqrt{|A_j|}} & v_i \in A_j \end{cases}$$

According to the properties of the Laplace matrix:

$$h_i^T L h_i = \frac{cut(A_i, \overline{A_\iota})}{|A_i|}$$

For subgraph *i*, its RatioCut is equivalent to $h_i^T L h_i$, so the goal of the algorithm turns to finding the smallest eigenvalue of L.

2.4.2.2 The Clustering Perspective

There is another type of algorithm in addition to the partition algorithm, for which constraint functions are used to identify the structure of the community. That is, instead of trying to find best cuts, these algorithms try to find the best aggregation scheme for combining the nodes in a graph into communities. This principle leads to a community detection algorithm in which the proximity between nodes is described by the presence and weight of edges between them. A branch of agglomerative algorithms in the above-mentioned hierarchical clustering exemplifies this approach. Aside from agglomerative algorithms, one of the most influential examples is Newman-Girvan modularity, one of the most common clustering metrics in the literature (Chen et al., 2014, Despalatović et al., 2014). Modularity is a global mass function that aims to find community structure from a global, whole network perspective.

The modularity measure is defined as (Newman and Girvan, 2004):

$$Q = \frac{1}{2m}\sum\nolimits_{ij}(A_{ij} - P_{ij})\delta(C_i, C_j)$$

Where *m* is the total number of edges of the graph, A is the adjacency matrix. If vertex *i* is connected with vertex *j*, $A_{ij} = 1$, otherwise $A_{ij} = 0$. P_{ij} is the expected number of edges between vertices *i* and *j* in the null model. $\delta(c_i, c_j)$ is a conditional function, if c_i=c_j, $\delta(c_i, c_j)$=1, otherwise $\delta(c_i, c_j)$=0. The goal of modular-based algorithms is to maximize the value of *Q*. If the number of within-community edges is no better than random, *Q* = 0. Values approaching *Q* = 1 indicate networks with strong community structure. In practice, values for such networks typically fall in the range of about 0.3 to 0.7. Higher values are rare (Wang et al., 2015). There are many modularity-based algorithms, including extreme optimization, spectral optimization, greedy optimization, simulated annealing, and genetic algorithms (Javed et al., 2018). However, all modularity-based algorithms have a common aim to place points in different communities to maximize modularity. Taking the Louvain method as an example, it uses a Greedy optimization that defines modularity as (Blondel et al., 2008):

$$Q = \frac{1}{2m}\sum_{ij}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j)$$

Where:

A_{ij} represents the edge weight between nodes *i* and *j*;

k_i and k_j are the sum of the weights of the edges attached to nodes *i* and *j*, respectively;

m is the sum of all of the edge weights in the graph;

c_i and c_j are the communities of the nodes;

If c_i=c_j, $\delta(c_i, c_j)$=1, otherwise $\delta(c_i, c_j)$=0.

Thus, the modularity value of community C is:

$$Q_c = \frac{\sum in}{2m} - (\frac{\sum tot}{2m})^2$$

$\sum in$ is the sum of edge weights between nodes within the community c.

$\sum tot$ is the sum of all edge weights of nodes within the community.

In the first step, each node is assigned a community. If removing a node from the community and joining it in another community cannot lead to a modular change, the node will be kept in the original community. Otherwise, it will be joined into the new community that caused the most significant modular increase. When all attempts have been made, the second step will be performed. The concept of hierarchy is introduced in the second step, whereby the community in the first step is used as a node to build a new network, and the first step is re-executed.

2.4.2.3 The Dynamical Perspective

These algorithms do not refer to algorithms developed for dynamic communities mentioned by Mittal and Bhatia (2021), but to algorithms that simultaneously consider the topology of the network and dynamic processes taking place within networks (Fortunato and Newman, 2022). As Rosvall et al. (2019) pointed out, for real-world networks (such as aviation networks), the structure of the network is naturally important, but understanding how the structure of the network affects the system's behaviour is also essential. Treating partitioning as a dynamic process thus distinguishes another type of community detection algorithm. Infomap (Rosvall et al., 2009) and Walktrap (Pons and Latapy, 2005) are two popular representatives. Both algorithms assume that a random walker is exploring the real network. Since the connections within a given community are closer, the random walker should be trapped in the community for a longer time. In other words, moving within a community is easier and moving between communities is harder.

The basic idea of the Walktrap algorithm is that when a random walker moves from one node to another with a given probability, short-distance random walks are more likely to remain within the same community. This characteristic is used to identify communities within a network (Pons and Latapy, 2005).

In the Infomap algorithm, the route of the random walker is recorded and used to identify communities. Communities receive unique codes based on the module switch rates of the random walker, while nodes within each community are encoded with the average node visit frequencies of an infinite random walk. When random walkers move from one community to another, Infomap records a unique exit code for the original community. Therefore, the path of a random walker in a community starts with the community code, ends with the exit code, and in the middle is the node code. When the starting point is within a community, other nodes in the same community are more likely to be visited, resulting in their codes appearing more frequently in the path. Consequently, accurately identifying and encoding communities can effectively reduce the code length of nodes within the network, thereby compressing the overall path information (see Figure 2.6). The cost function of the Infomap algorithm is the length of the code record, or in other words, the information cost for describing the movements of the random walker. The shorter the length, the better the algorithm's performance (Rosvall et al., 2009).

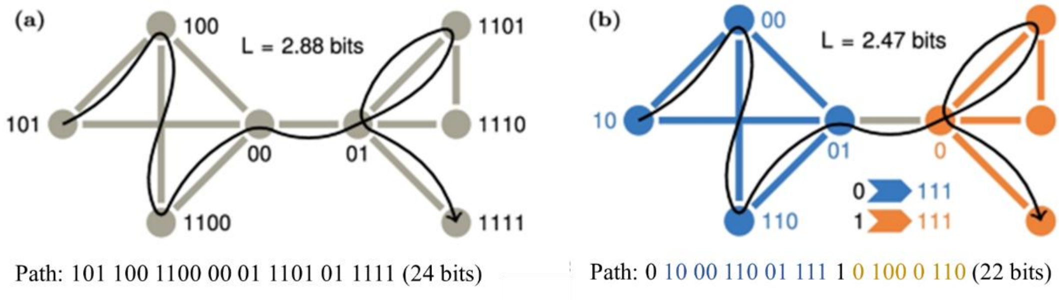


Figure 2.6 Example of how community delineation affects path information length. Different colours represent different communities. The numbers to the left and right of the arrows represent the entry and exit codewords for each community, respectively. The average per-step code length is denoted as L (Blöcker et al., 2022).

Given a partition *M*, the description length *L(M)* that defines the average walk path is (Rosvall et al., 2009):

$$L(M) = q_{\curvearrowright} H(\zeta) + \sum_{i=1}^{m} p_{\circlearrowright}^{i} H(\mathrm{P}^{i})$$

$q_{\curvearrowright}$ is the probability of going from one module to another in a random walk step; $H(\zeta)$ is the entropy of the module encoding; $H(\mathrm{P}^{i})$ is the entropy of the encoding of the node in the walk, measuring the average information required to describe the walker's steps; $p_{\circlearrowright}^{i}$is the probability that module i is visited.The calculation process of Infomap is similar to Louvain's algorithm, except that the modularity is replaced by $L(M)$.

2.4.3 Applications of algorithms in water pipeline network research

The application of community detection in the study of water networks lies mainly in the delineation of District Metering Areas (DMAs). The aim is to use community detection algorithms to identify areas that can be isolated for pressure management and leakage reduction. However, DMA delineation must incorporate the physical and hydraulic characteristics of the network such as topography, water pressure, and water quality (Khoa Bui et al., 2020, Sharma et al., 2022b). Consequently, studies tend to optimise community detection algorithms based on these properties, either manually or automatically (Khoa Bui et al., 2020). A summary of examples using community detection algorithms can be found in Table 2.5. Additionally, given the distributional similarities of road and water networks, and the extensive research in road networks, applications of community detection in road network studies are also listed here for reference.

Table 2.5 Summary of objectives and algorithms used in previous community detection studies of piped or road networks via network analysis (DMA: District Metered Area)

Reference	Algorithm	Application Areas	Objective
(Brentan et al., 2017)	Walktrap	water	DMA creation
(Brentan et al., 2018)	Walktrap	water	DMA creation
(Campbell et al., 2014)	Walktrap	water	DMA creation
(Jazayeri and Moeini, 2024)	Improved Girvan-Newman algorithm	water	DMA creation
(Scibetta et al., 2013)	fast-greeding modularity	water	DMA creation
(Sharma et al., 2022a)	fast Newman algorithm	water	DMA creation
(Zhang et al., 2017)	Louvain	water	DMA creation
(Haghbayan et al., 2021)	Infomap	traffic	congested urban road identification
(Hong and Yao, 2019)	Infomap	traffic/urban areas	functional area identification
(Shang et al., 2020)	Order Statistics Local Optimization Method	traffic	assessment of the urban road network
(Bramson, 2022)	greedy modularity communities algorithm	traffic/urban areas	neighbourhood identification

(Law et al., 2019)	modularity optimisation algorithm	traffic/urban areas	neighbourhood identification

Several review articles also describe the application scenarios for different algorithms and explain their conceptual advantages and disadvantages. Gates et al. (2016) evaluated weighted, undirected community detection algorithms from the perspective of applications in brain science. They evaluated the Spectral Approach, Walktrap, Fast Modularity, Louvain method, Label Propagation and Infomap algorithms. The results indicate that when the data are in the form of sparse count networks (such as those seen in diffusion tensor imaging), Label Propagation and Walktrap surfaced as the most reliable methods for community detection. For dense, weighted networks such as correlation matrices capturing functional connectivity, Walktrap consistently outperformed the other approaches for recovering communities. Wickramasinghe and Muthukumarana (2022) also compared the performance of the algorithm in both sparse and dense networks, and found that the Louvain algorithm performed well in both contexts. Harenberg et al. (2014) compared the performance of eight algorithms across five networks and concluded that an algorithm's output with a good community structure does not necessarily have high accuracy, and vice versa. Smith et al. (2020) described application scenarios for the Edge-Betweenness, Random Walktrap, Label Propagation, Infomap, Louvain, and Spinglass algorithms. Their study suggests that researchers should choose an algorithm based on the main research problem in conjunction with the principles of the algorithm. Ghasemian et al. (2019) compare the performance of 11 algorithms for a specific task, indicating that algorithms sharing similar underlying assumptions tend to exhibit comparable performance, although the similiarity remains contingent upon the characteristics of the network. Fortunato and Hric (2016) mentioned in their algorithm guide that the methods based on modularity and Infomap algorithms perform better. Infomap is more easily adapted to different types of input data and research questions. In contrast, the modularity algorithm itself has a resolution problem.

Particularly, graph studies on water networks face significant constraints due to limited data availability. Although a systematic review on this topic is absent, it is noteworthy that several pipeline studies in the table used synthetic rather than real-world networks. This exemplifies a broader trend in water research, as seen in the 44 publicly available datasets of water pipeline networks listed by Giudicianni et al. (2018), of which 23 are synthetic. Momeni et al. (2023) highlighted that the paucity of real-world data assets is one of the primary obstacles to research on water distribution networks. Furthermore, Yu et al. (2024) emphasized that this limitation not only restricts the study of network topology but also hampers the transferability of findings to different regions.

In summary, on the one hand, the performance of algorithms differs across various scenarios; on the other hand, experience in applying algorithms is lacking in some regions. Thus, although community detection has been applied to road networks for urban planning and used in water network studies, there remains a research gap in understanding the relationship between water infrastructure and urbanisation trajectories, particularly in slums.

2.5 Water access in urban areas: performance and measurement

SDG 6.1 aims to achieve universal and equitable access to safe and affordable drinking water for all by 2030. The WHO/UNICEF Joint Monitoring Programme (JMP) has developed a classification system for drinking water facilities and services to benchmark and monitor progress, with water access being a core component of this system (WHO/UNICEF, 2023b). There are significant disparities in water access between regions—beyond the impacts of economic and social factors, the availability of improved water sources varies between countries and cities (Deshpande et al., 2020, Dos Santos et al., 2017, Wagle, 2022), urban poor are confined to slums or peripheral areas and live far from the areas where these facilities are concentrated (Armah et al., 2018). To better understand how regional characteristics influence water supply and to examine water access patterns in informal settlements, researchers have called for spatial studies of water access (Cassivi et al., 2019, Dos Santos et al., 2017).

2.5.1 Place-based approaches to measuring water access

Many measures have been developed to study infrastructure accessibility. Generally, accessibility refers to the ease of reaching urban services (in this case, water) or the interaction between people and infrastructure (Chen et al., 2017). According to Siddiq and Taylor (2021), the four types of factors that can influence accessibility include the built environment, transport systems, individual characteristics, and trip characteristics (see Figure 2.7). Studies define accessibility differently and focus more on some of these four factors. Their different methodologies can be categorised into four groups: infrastructure-based measures, location-based measures, person-based measures, and utility-based measures. They measure infrastructure accessibility in terms of facility performance, spatial distribution of facilities and population, individual activities, and economic benefits, respectively (Geurs and van Wee, 2004, Higgins et al., 2022). To simplify, we can classify them into place-based accessibility measures, analysing the spatial proximity to urban opportunities, and person-based accessibility measures, measuring the spatial and temporal constraints individuals experience when travelling (Chen et al., 2017).

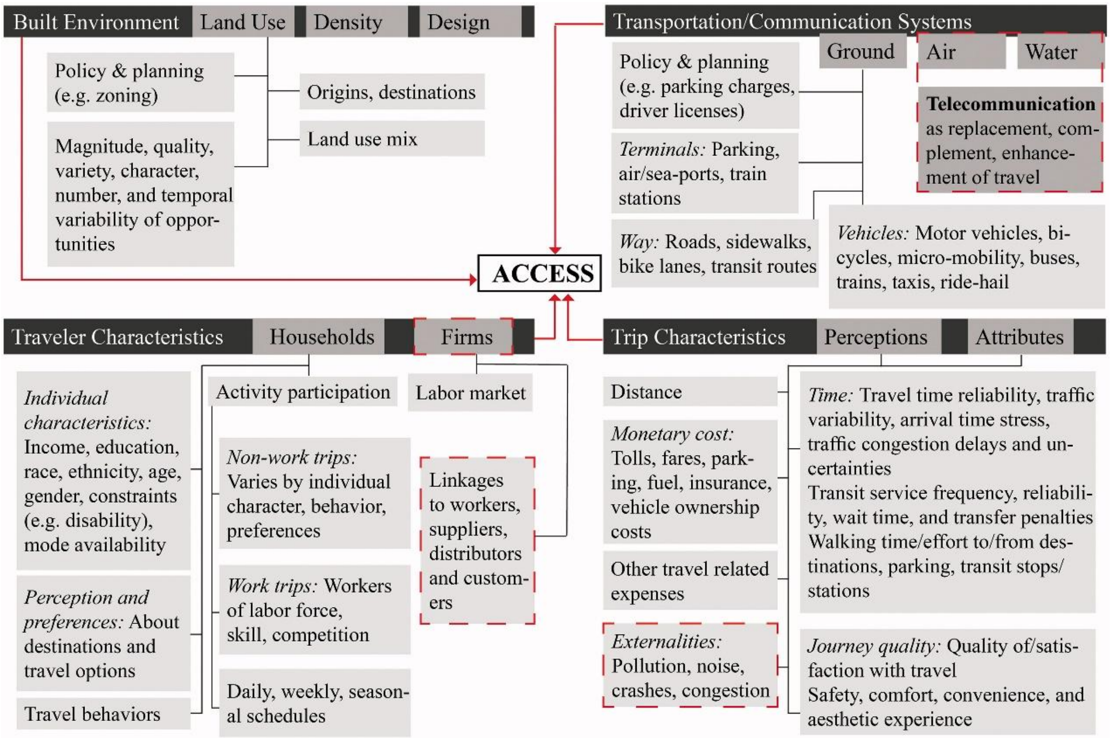


Figure 2.7 A conceptual model of the factors affecting accessibility (Siddiq and Taylor, 2021).

More specifically, person-based accessibility measures emphasise the impact of travellers' characteristics on accessibility. Such a concept is derived from time geography, which suggests that accessibility is constrained by the individual's daily activity schedule (the time budget) and the travel time to the location of the target activity (Liao, 2019). Unsurprisingly, while it is possible to measure differences in accessibility between individuals and provide insight into how personal traits interact with the environment and affect accessibility, person-based methods rely on detailed personal information and travel logs (Boschmann and Kwan, 2008) (Huang, 2019). For example, travel diaries were used in the studies by Neutens et al. (2012) and Dixit and Sivakumar (2020). Understanding person-based accessibility patterns also relies on people's perceptions, the domain of perceived accessibility, which requires more comprehensive and flexible models (Bugden and Stedman, 2019, Pot et al., 2021). Firstly, studies that emphasise the importance of personal experience also acknowledge the influence of spatial elements on accessibility. Secondly, as demonstrated by Siddiq and Taylor (2021) and Miller (2018), scholars often expect to build a comprehensive model to include the effects of individual/household characteristics on accessibility. However, such models are often flexible, complex, and difficult to interpret, and the variability of individual behaviours can make person-based availability approaches both challenges and difficult to generalise.

Instead, more operationally pragmatic models in practice come from place-based measures (Li et al., 2023a). These approaches focus more on the other three elements in addition to individual characteristics (Luo and Qi, 2009):

1. Supply, the amount, capacity and distribution of the facilities, and the quality of the service.
2. Demand, the demand for water by population groups (of different ages, genders, incomes, etc.) in the study area, and their distribution pattern.
3. Spatial distance, distance between population and water sources, travel modes and corresponding travel speeds, the negative effect of travel time on water accessibility, and competitive relationship between facilities.

These three elements form easy-to-grasp accessibility measures. The two most popular types are cumulative-opportunity and gravity-based approaches (Bittencourt and Giannotti, 2023). Given a threshold, cumulative opportunity measures count the number of opportunities that can be reached by the threshold time, and opportunities beyond the threshold are not considered (Kelobonye et al., 2020). The strong relationship between thresholds and cumulative accessibility results in a major drawback of traditional person-based approaches. Recently, a time interval cumulative accessibility measure using multiple thresholds has been proposed to overcome the shortcomings of the traditional cumulative opportunity measure (Tomasiello et al., 2023). Nevertheless, it is still a simple and direct accessibility model, while gravity-based approaches are considered more complex. In the gravity model (Weibull, 1976), the accessibility of a population location depends on the supply/demand ratio of all facilities to the population within a certain area and is adjusted according to the distance between them. More specifically, the supply/demand ratio of a facility to all population within its catchment is the attractiveness or "opportunity" of the facility. Facility attractiveness decreases as distance to a given population increases. All opportunities a population location can approach collectively constitute the location's accessibility.

As highlighted by Miller (2018), there is no objective definition of the accessibility score, thus accessibility score only has meaning when being compared within a group. So, deciding between gravity-based and cumulative opportunities methods is difficult, especially when their performances are similar. For instance, studies using cumulative opportunity and gravity-based approaches in London and Vancouver show that they have similar results given the same thresholds (Kapatsila et al., 2023). However, a recent study points to a potential drawback of cumulative-opportunity approaches. Klar et al. (2023) used cumulative-opportunity, gravity-based, and a hybrid of the two approaches for Vancouver's transit system. They observed that since the cumulative-opportunity approach only considers opportunities within a given catchment and innovations often lead to changes in service coverage areas (catchments), gravity-based models respond better when examining the range of impacts.

2.5.2 The two-step floating catchment technique: methodological developments and applications

The two-step floating catchment area (2SFCA) method, which has been widely used in recent years, is derived from the gravity-based model. The problem with the traditional gravity model is that, while accounting for the negative effect of distance on attractiveness, the model does not define the catchment of the facility and population location. This results in the number of facilities that can be accessed by the population location as well as the accessibility score being sensitive to study area delineation and choice of administrative geography, thus making it hard to interpret the outputs (Luo and Qi, 2009). As a special case of the gravity model, 2SFCA corrects this problem by defining the catchments and allowing the boundaries to float based on a user-specified distance from each facility and population location, thus limiting the spatial variables that are used in the calculation. Therefore, the 2SFCA method can be implemented in two steps:

Step 1: For facilities *j*, search all population locations *k* that are within a threshold travel time from location *j*, and compute the capacity-to-population ratio R_j, defined as the supply S_j e.g., facility capacity) divided by the population within the catchment area *P*, discounted by distance decay function $f(d_{kj})$.

$$R_j = \frac{S_j}{\sum_{k\in\{d_{kj}\le C_j\}} P_k f(d_{kj})}$$

Step 2: For each population location *i*, search all facilities locations (*j*) that are within the catchment area of population location *i* (C_i), and sum up the capacity-to-population ratios (derived in step 1), R_j, at these locations, discounted by distance decay function $f(d_{ij})$:

$$A_i^F = \sum\nolimits_{j\in\{d_{ij}\le C_i\}} R_j f(d_{ij})$$

In recent years, several improvements have also been introduced to the 2SFCA model, one of which focuses on the distance decay function that measures the negative effect of distance on accessibility. In the classic 2SFCA model, dichotomous methods are used, which divide the catchment into zones with different weights based on distance (Liu et al., 2022). In other words, the classic methods simulate the negative effect of distance on accessibility by directly adding or subtracting weights for each zone. This leads to a similar situation as opportunity accumulation models: two points inside and outside a boundary will have very different weights, and different locations in the same zone will have the same weight, which does not align with common sense. Therefore, continuous distance-decay functions were introduced to simulate travel friction. They imply that distance is negatively correlated with accessibility. As the distance increases, the weights gradually decrease. Commonly used functions include the exponential, Gaussian, and kernel density functions. Among them, the Gaussian function is believed to be the best-performing function in access studies (Chen and Jia, 2019, Liu et al., 2022). It is defined as:

$$f(d_{ij}, d_0) = \frac{e^{-(1/2)} \times (d_{ij}/d_0)^2 - e^{-(1/2)}}{1 - e^{-(1/2)}}$$

Another improvement on the original 2SFCA model is the realisation that in addition to the capacity of the facility, the demand also needs to be weighted according to the crowdedness of facilities. In other words, there is a need to incorporate people's preferences for facilities, which are considered to be related to distance (Kanuganti et al., 2016). In practice, when there are multiple facilities that people can access within a threshold time, this model enhancement assumes that the closest one with the best service quality will be chosen. Hence, residents in a population location will make decisions about attending services based on trip distances and capacity at service points. It also means that a facility within a population catchment will not be accessed by all residents of that population location. However, according to the 2SFCA formula above, each facility's supply/demand ratio is calculated using the total head count of a population location, which means that the traditional 2SFCA overestimates demand (Subal et al., 2021). Therefore, there is need for an indicator to measure facility crowdedness or offset the overestimated demand. Given this need, the Huff function has been used to improve the 2SFCA (Wang, 2018). The Huff function searches for all facilities *k* that a population centroid *i* can approach within the travelling time threshold and compares the performance of a particular facility *j* with the rest of the facilities on the distance decay function. If facility *j* is closer to population centroid i than other facilities and has more capacity, then facility *j* will have a higher Huff value than the other facilities. It will also have a higher likelihood of becoming the destination of population centroid, *i*.

$$Prob_{ij} = \frac{S_j f(d_{ij})}{\sum_{k \in [d_{ik} \le d_0]} S_k f(d_{ik})}$$

Where $Prob_{ij}$ is the probability of *i* choosing *j*.

Based on these two improvements, the improved 2SFCA method can be represented as:

Step1:

$$R_j = \frac{S_j}{\sum_{k \in [d_{kj} \leq d_0]} P_k f(d_{kj}) Prob_{kj}}$$

Step 2:

$$A_i^F = \sum_{j \in [d_{ij} \leq d_0]} Prob_{ij} f(d_{ij}) R_j$$

2SFCA and improved versions have been widely used in studies on public facility accessibility, especially in healthcare studies (Kanuganti et al., 2016, Luo and Qi, 2009, Wang, 2012). The experience in these healthcare accessibility studies shows that 2SFCA is sensitive to parameter selection, especially in the choice of distance thresholds. Chen and Jia (2019) pointed out that variants of the 2SFCA model yield similar results if the same threshold value is used. This proves that the threshold value is the main factor that affects the results. A large threshold may smooth the spatial pattern of the accessibility map, while a small threshold introduces more localised variations into the result. They also pointed out that since thresholds are also variables in the distance decay function, it will also affect the performance of the distance decay function. Luo and Whippo (2012) emphasised the significance of the threshold affecting the accessibility results. They argue that the distribution of facilities and population differs significantly between urban and rural areas. As a result, the time urban versus rural residents are willing to spend on accessing facilities varies, and using the same catchment sizes will overestimate accessibility in both areas. Therefore, when applying the 2SFCA method, thresholds should be chosen carefully with attention to the specific application and relevant theory.

As highlighted in previous discussion of global urbanism (see section 2.2.1), urbanisation and new interventions have made the landscape of basic services in these regions increasingly complex and fragmented (Coutard, 2008, Smiley, 2020). On the one hand, the emergence of informal settlements increases the complexity of urban spaces; on the other hand, non-piped water supply systems play a crucial role in water access in global southern cities (Adams, 2018b, Azunre et al., 2022). Both of these factors can be captured by the 2SFCA method. The results of this method are influenced by population and facility capacity, and the sensitivity of the 2SFCA method to parameters corresponds to the spatial heterogeneity of urban water access. Thresholds in 2SFCA models represent how much time residents are willing to spend accessing services and the travel modes (walking or driving) used by an area's inhabitants (Chen and Jia, 2019). This means that 2SFCA results can be interpreted in terms of the size of the thresholds: they reveal the distribution of facilities and populations, as well as the social factors associated with accessibility (Wan et al., 2012). Therefore, 2SFCA can be used to analyse the complex water supply landscape in cities of the Global South. However, while the 2SFCA method and its improved versions have been extensively applied in healthcare studies (Kanuganti et al., 2016, Khashoggi and Murad, 2021, Luo and Qi, 2009, Shao and Luo, 2022, Tao et al., 2020, Wang, 2012), its application in the water sector remains largely unexplored, despite repeated calls for more quantitative evidence on water access (Cassivi et al., 2019, Dos Santos et al., 2017).

Chapter 3 Research scope and technical resources

Rapid urbanisation in Africa poses significant challenges to land, infrastructure, and service provision in cities, further exacerbating urban inequality (UN-Habitat, 2022). Sections 2.1 and 2.2 have outlined how this urbanisation intensifies pressure on urban water services. According to the UN reports on the Sustainable Development Goals (SDGs) (UN-Habitat, 2023b, UN, 2023a), Sub-Saharan Africa (SSA) has some of the highest concentrations of slum populations and the largest numbers of people lacking access to basic water services. Various initiatives, such as slum upgrading and delegated management model (DMM), have been introduced to address these pressing issues, generating diverse responses within the context of urbanisation in southern cities.

Recent studies have proposed different theories to understand and guide these interventions, yet they remain constrained by the lack of empirical evidence (Parida and Agrawal, 2023, Pieterse, 2011). Particularly in SSA, the impact of urban space and water interventions on service delivery remains largely unknown and uncertain. While graph and access measurements have been shown to be effective in assessing infrastructure patterns and the urban environment, neither has been fully developed or applied specifically for water study purposes (Derudder and Neal, 2018). Since these methods often involve empirically adjusted parameters, this leaves a research gap.

To address the gaps and establish links between research methods and the theoretical frameworks of urbanisation and water services in SSA, this study seeks to gather empirical evidence from cities in SSA by analysing water infrastructure in southern cities from multiple perspectives. This chapter primarily outlines the selection of the study areas and provides relevant background information.

3.1 Case study cities: description and rationale

To achieve the research aim of analysing the impact of SSA urbanisation on water supply, the study area must ideally have rich data on water infrastructure, particularly pipeline data, as this is essential for applying graph theory and accessibility analyses. Additionally, the area should share similar urbanisation characteristics, specifically facing challenges of slum expansion and water access, to ensure the representativeness of the study and allow for broader applicability of the results across various SSA cities. Based on these criteria, this section reviews available databases and recent SDG monitoring reports, and provides a description of the urban layout, road networks, and water infrastructure of the two selected case study areas, Kisumu in Kenya and Kigali in Rwanda, both located in East Africa. This description lays the foundation for interpreting the subsequent analyses.

3.1.1 Rationale for selection of case study cities

An audit of water infrastructure data to assess the data landscape was the first step in selecting potential study areas from SSA countries. Since the main providers of water services in cities within SSA are usually governmental or regional/international NGOs/humanitarian organisations, this audit entailed structured searches of international and regional databases for water infrastructure and consumer/supplier data. The databases searched included the World Bank, International Benchmarking Networking, Food and Agriculture Organization, World Resources Institute, Water Point Data Exchange, OpenStreetMap, openAFRICA, and Africa Infrastructure Knowledge. For point data, tags related to water infrastructure (e.g. water treatment plant, pumping station, water tank) were of interest.

Following the searches, it became apparent that very few open data sources cover SSA. Datasets such as the World Bank and Humanitarian Data Exchange only provide water statistics and analysis reports for water services. At the same time, the only databases containing geospatial infrastructure data are OpenStreetMap (OSM) and Water Point Data Exchange. A search of OSM showed that, as of 2022, there were only 1,539 pipeline records in Africa, mainly in South Africa, Libya, and Ethiopia. Since Libya is not an SSA country and Ethiopia's records are mainly distributed in the rural areas, only South Africa was investigated further. There are 161 records in South Africa, but in many cases, a single pipeline may be associated with multiple records, meaning the actual number of features is lower.

In comparison, there is a greater availability of point data related to pipelines. The Water Point Exchange database has 14,434 records for piped water points in SSA countries (as of 2024), with 11,385 of those being public tapstands and kiosks. Sierra Leone and Ghana have the highest number of water point records. Despite the fact that many searched tags are not available in Africa, the survey of OSM data shows that there are 18,923 large water points (where larger amounts of drinking water can be collected) and 21,144 general drinking water in Africa. These points are mainly located in Uganda, Kenya, and Burkina Faso. The low number of pipeline records compared to the high number of ancillary facilities suggests that these SSA countries have dense piped infrastructures, and their pipeline networks are not yet covered by public domain databases. These counties include Sierra Leone, Ghana, Liberia, Zambia, Kenya, Uganda, and Rwanda.

Another key finding from the data audit is that the data quality of public databases such as OSM and Water Point Data Exchange varies across SSA, largely due to the type of data source. For example, Sierra Leone has well-documented water points due to government efforts, while water point clusters in Liberia are associated with the Liberia Firestone Rubber Company. In some instances, OSM, as a Volunteered Geographic Information (VGI) platform, can be incomplete with inconsistent labelling, particularly in Africa. In these cases, utility companies with established GIS-based asset management systems are the most reliable data sources.

After further investigations, two cities, Kisumu and Kigali, were further filtered from the above countries as the study areas for this research. Both cities face rapid urbanization and high proportions of slum populations: slums in Kisumu account for 19% of the city’s area and house 60% of its urban population (Othoo et al., 2020), whereas 77.3% of Kigali’s households live in slums, covering 79% of the population (Hitayezu et al., 2018, NISR, 2018). Despite the similarities, their piped water service levels differ significantly. In Kisumu, Kisumu Water and Sanitation Company (KIWASCO) provides piped water to 8.3% of households within dwellings, 10.9% to yards/plots, and 22.1% via public taps (KNBS, 2019b). Water coverage in Kisumu’s low-income areas has increased, but unregulated connections remain, creating unsafe "spaghetti" networks (Boakye-Ansah et al., 2019, LVSWSB, 2021). In contrast, Kigali’s Water and Sanitation Corporation (WASAC) reported a 50% household water connection rate in 2018, with about 87% of unplanned settlement households accessing improved water sources (City of Kigali, 2020a, Hitayezu et al., 2018, NISR, 2018). Additionally, Kisumu was the first city in Kenya to implement a large-scale water DMM program targeting slums (Nzengya, 2015). By 2022, KIWASCO had partnered with 41 master operators to manage water delivery in low-income areas under DMM contracts. Unlike Kigali, where WASAC directly manages all water infrastructure, Kisumu’s approach includes some decentralized management to address water accessibility challenges. This contrast provides an opportunity to compare their existing pipeline networks and explore the interactions between SDG 6 and SDG 11, particularly in terms of policy impacts.

3.1.2 Urbanisation and infrastructure development in Kisumu

3.1.2.1 Urban layout

Kisumu County is located in western Kenya. The topography of Kisumu County varies, with northern regions characterized by hilly terrain and southern regions predominantly consisting of plains. The county is bordered by Lake Victoria, the world’s second-largest freshwater lake, to the west, and by mountains to the east. Within this, Kisumu County has a total area of 2,085.9 km^2. According to the 2019 census, Kisumu County has a total population of 1,155,574, with 397,957 located in Kisumu City, the third-largest city in Kenya after the capital Nairobi and Mombasa (KNBS, 2019a).

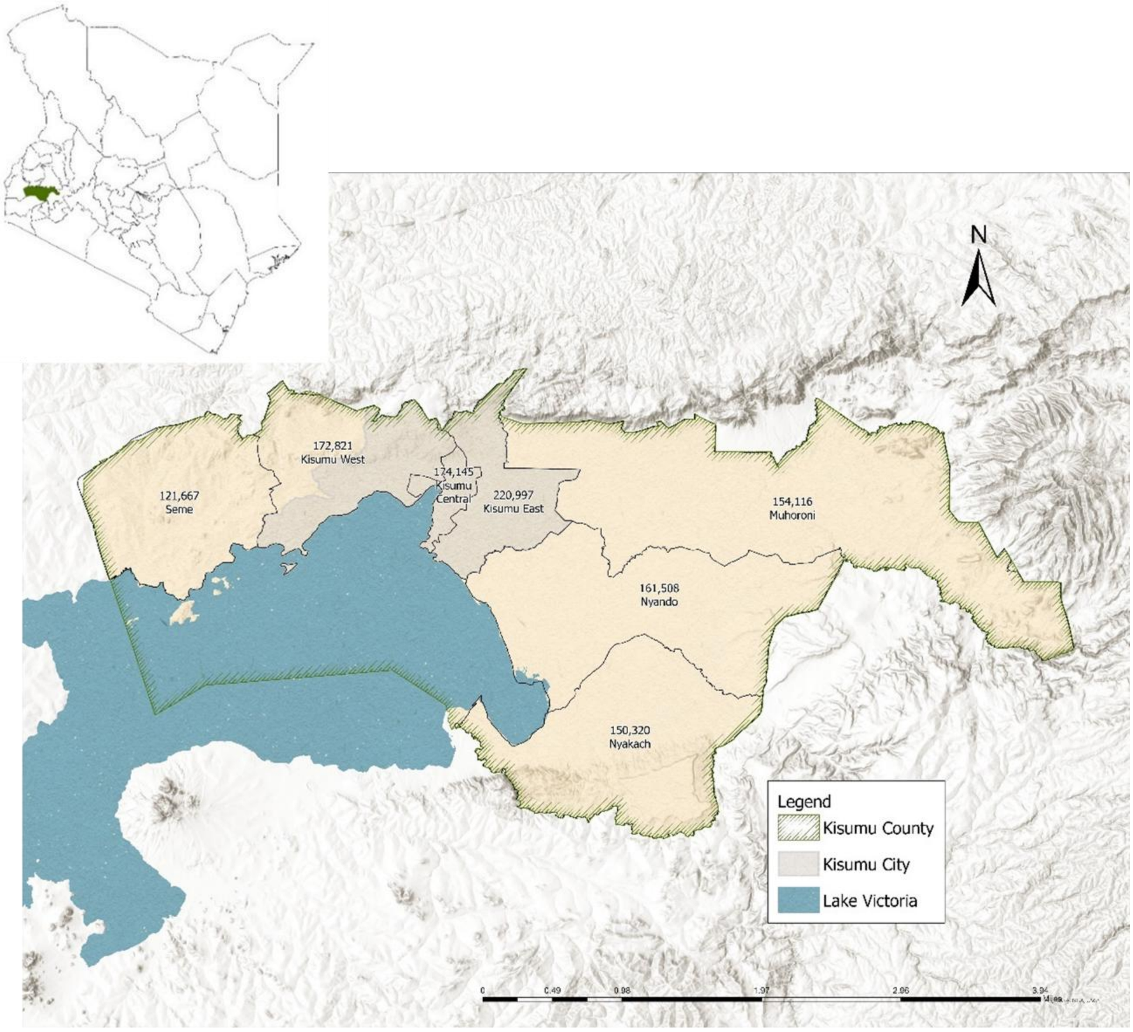


Figure 3.1 Map of Kisumu County, showing its constituent 2019 municipal sublocations including Kisumu City. Names and population counts from the 2019 Kenya census (KNBS, 2019a) are labelled per each sub-county. The overview map shows Kenya's 47 counties with Kisumu marked in green.

The name Kisumu stands for the trade of goods, which comes from Kisumu's longstanding status as a commercial and transportation centre in the Lake Basin region. In 1898, the tip of the Winam Gulf was designated as the railway terminus for the Kenya-Uganda railway. With this as the central area, Kisumu City gradually developed into its current size. The name Kisumu can refer to two entities: Kisumu County and Kisumu City. Kisumu County is one of Kenya's 47 counties and is subdivided into 7 sub-counties (as shown in Figure 3.1). However, 'Kisumu' is more frequently used to refer to Kisumu City, the study area for this study. The boundaries of Kisumu City are largely inherited from the previous Kisumu Municipality. Today, Kisumu City is divided into 25 sub-locations or 10 main areas (Township, East Kolwa, Central Kolwa, Southwest Kisumu, North Kisumu, Central Kisumu, East Kisumu, West Kajulu, East Kajulu, and West Kolwa) (Figure 3.2). From the perspective of history and planning, those sub-locations can be grouped as central urban areas and informal areas, representing the developed areas in central Kisumu and less developed areas in peripheral regions, respectively.

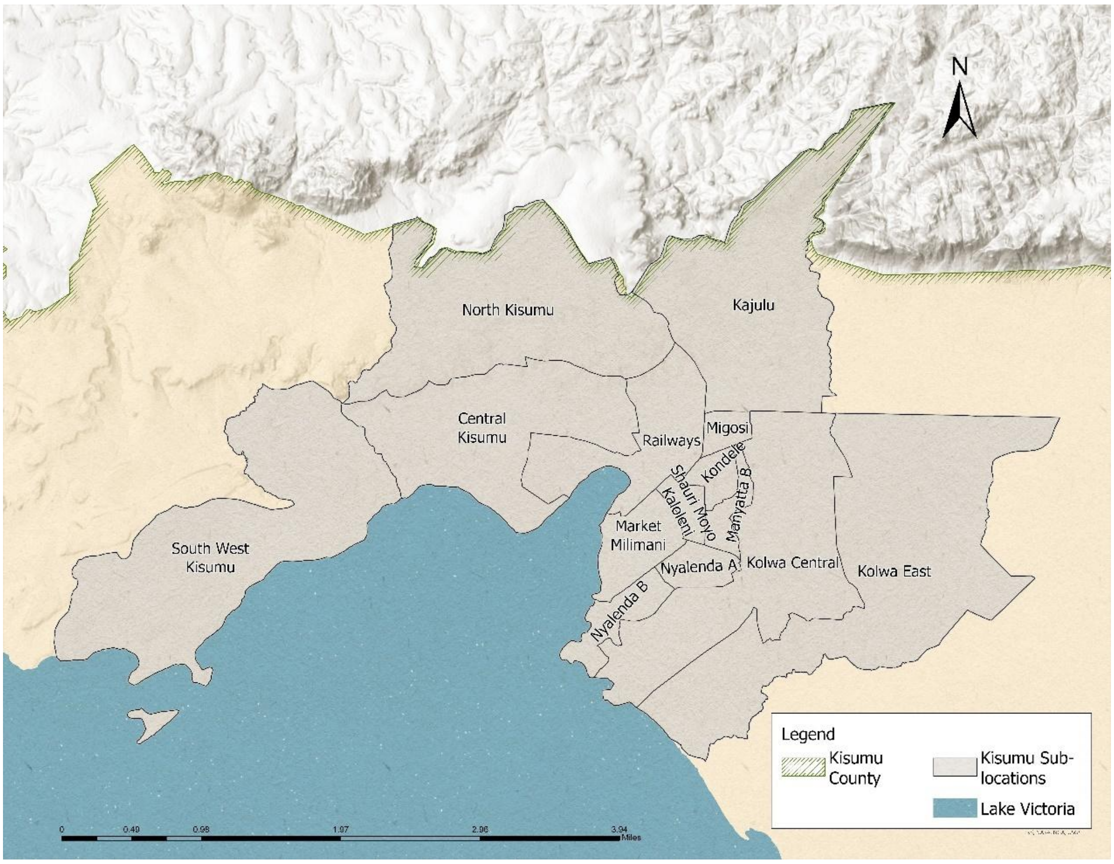


Figure 3.2 10 main administrative areas of Kisumu City.

The central urban areas are the well-planned neighbourhoods for the middle and high-income residents, including the former European residential area (Milimani), the former Indian residential area (Kibuye), low and middle-income public housing (Municipal, Railways, Kenya Post, Kenya Power). These areas originated from colonial-era plans and were characterised by high levels of service delivery and infrastructure provision (Letema et al., 2014, UN-Habitat, 2005). Apart from colonial planning, a second reason for the distinction between the central urban areas and the informal areas was the extension of the city's boundaries in 1972. This plan increased the city's total area to 53 km^2. The peri-urban region that was included now fell under the administration of Kisumu County Council whereas previously, it fell under the rural Kisumu District administration. This means that the planning policy historically applied to the newly extended area was the rural standard. Since the rural standard is lower than the urban standard, there has been a persistent difference between the informal and central areas since then, resulting in two very different urban configurations within Kisumu (UN-Habitat, 2005).

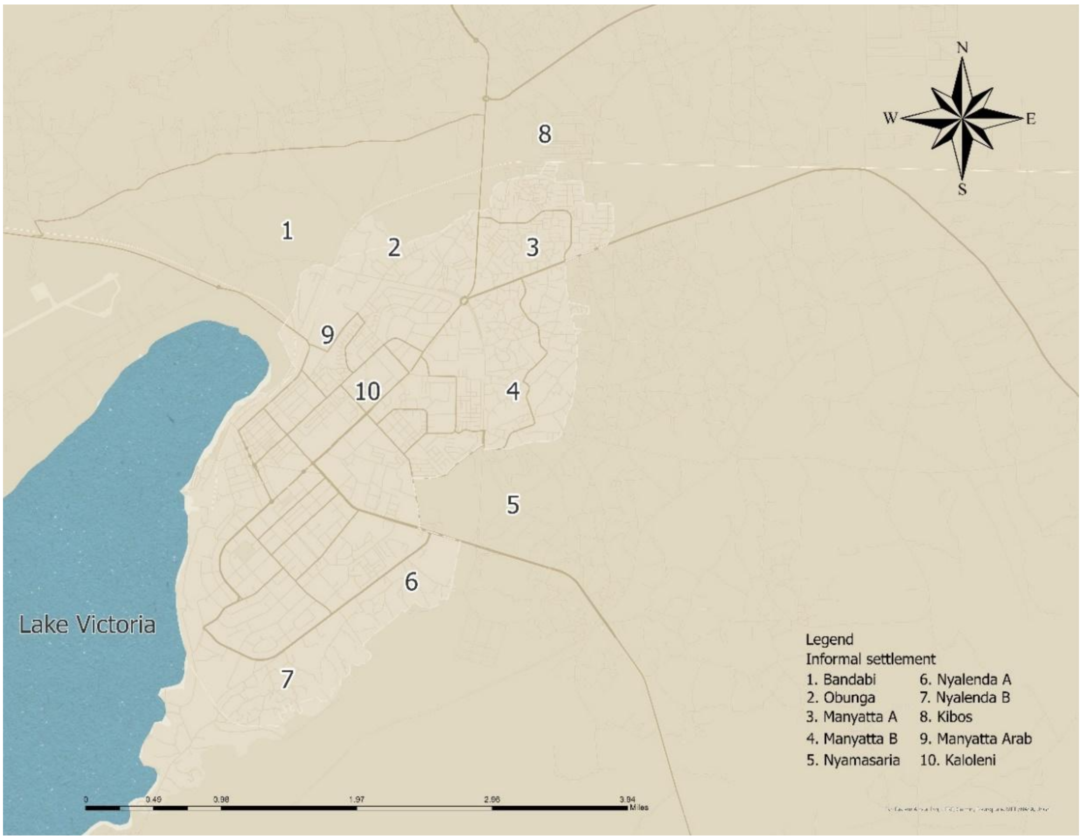


Figure 3.3 The slum belt surrounding the central areas in Kisumu (UN-Habitat, 2005).

As shown by Figure 3.3, unplanned informal areas surround the central areas of Kisumu. As a result of colonial zoning, African communities were isolated from the city's core and housing developed without regulation and planning (Letema et al., 2014). In recent years, many immigrants from other counties have also settled here for employment and entrepreneurial opportunities. These slums cover approximately 19% of the city and host about 60% of Kisumu's urban population (Othoo et al., 2020). These slums and central urban areas form the boundaries of Kisumu City, beyond which are the vast rural areas and satellite towns of Kisumu County.

Kisumu City has a higher population but smaller area than the remaining areal units. The urban population of Kisumu County is 1,155,574, while the Kisumu East (220,997), Kisumu West (172,821), and Kisumu Central (174,145) are the three sub-counties with the largest populations in Kisumu County. Due to boundary changes since the last census in Kisumu (in 2009), it is hard to analyse the spatial distribution and growth rate of the population in more detail. Nevertheless, it can be concluded that these slums, concentrated in high-population sub-counties, are home to a significant proportion of Kisumu's population, which increases pressure on infrastructure. The 2019 Kenya Population and Housing Census Volume IV also notes that the proportion of population with piped water into dwellings in Kisumu has an extreme distribution. The average proportion in Kisumu County is 8.3%, with the highest proportion in

Kisumu Central, at 26.1%. The Seme has the smallest proportion, 0.7%. And for sub-counties other than Kisumu Central, the figure is no more than 10% (KNBS, 2019b).

3.1.2.2 Water services

Since the 2010 Kenyan constitution made access to safe water a human right, subsequent laws have divided powers and responsibilities for water and sanitation services (Price et al., 2018). Each county provides services within its jurisdiction while being regulated by the state. In Kisumu County, the Municipal Council of Kisumu (MCK) owns all water and sewerage facilities, and institutions such as the Water Resources Authority are responsible for managing water resources. Since 2003, the Lake Victoria South Water Services Board (LVSWSB) has been responsible for executing and implementing water projects and licensing water service providers. The actual water service provider is Kisumu Water and Sanitation Company (KIWASCO).

Most of the water in Kisumu County is obtained from Lake Victoria, but some residents also rely on rivers (such as Kibos, Nyamasaria, Kisian, Kajulu, Mamboleo, Luanda, and Lidango) and groundwater (Maoulidi, 2010). In central urban areas, the pipelines are mainly managed by KIWASCO, and their spatial distribution is regular, mainly laid below the streets. Statistics from the Kisumu County Water Resource Masterplan Draft Report show that in low-income areas (peri-urban or informal areas), the coverage of the water supply network is still very low (LVSWSB, 2021). In the absence of supervision, operators in informal areas tend to connect pipelines to the network in illegal and unplanned ways. The structure of these pipelines is called a 'spaghetti' network. In addition to inefficient spatial configuration, 'spaghetti' lines also face safety and quality issues (Boakye-Ansah et al., 2019).

It is important to consider Kisumu’s water service development within the broader historical context of Kenya. Following the economic crises and rapid urban population growth from the 1970s onwards, Kenya’s water sector faced severe financial deficits, prompting a series of reforms. These reforms included the introduction of local authorities and the adoption of commercial models to manage and operate water infrastructure (Nilsson and Nyanchaga, 2008). This shift in policy aligns with the context discussed in Section 2.2 of the literature review, where the emergence of public-private partnerships (PPP) and slum upgrading was emphasised. In Kisumu, water service provision and development were commercialised under the Kenya Water Act 2002, which provided the legal framework for the introduction of DMM (GoK, 2002). In Kisumu, KIWASCO is responsible for providing water in bulk to agents (master-operators), who in turn deliver water to consumers through retail or by establishing pipelines (Nzengya, 2015). KIWASCO will provide them with basic network extension and maintenance training. However, in practice, those master-operators tend to build pipeline networks to a lower standard for reasons of limited space in slums and cost saving, resulting in poorer-quality pipeline and water distribution (Nzengya, 2015).

According to the Kisumu County Water Resource Master plan Draft Report (LVSWSB, 2021), the challenges faced by Kisumu's water and sanitation services are:

1. Growing demand for services. The water demand in 2020 was 78,332 m^3 /d. Considering the development of population and income, water demand is expected to be 87,248 m^3 /d in 2025, and by 2050, this amount will double.
2. The infrastructure needs to be renewed and expanded. Existing water pipelines suffer from deteriorating condition and poor coverage. The previous stage (the second phase of the Kisumu Water Supply and Sanitation project) of Kisumu's infrastructure development scheme mainly focused on water intake and treatment works, transmissions mains and storage tanks and ignored network maintenance and expansion of the secondary distribution system. Therefore, existing pipelines are often at risk of blockages and leaks. In areas already covered by the network, the capacity of the network lags behind demand. In low-income areas such as informal areas, water services coverage remains very low, and even current needs cannot be met.
3. The DMM approach requires improvement and expansion.
4. Data to support network renewal and regular maintenance is lacking.

3.1.2.3 Roads

The road network in Kisumu is managed by three agencies: the Kenya National Highways Authority (KeNHA), the Kenya Urban Roads Authority (KURA), and the County Government of Kisumu (Figure 3.4). Due to Kisumu's location, the main arteries serve dual roles: they are both urban streets and crucial transportation routes connecting other cities. This dual function is one of the factors influencing road design in Kisumu (County of Kisumu, 2020). Interventions have been implemented to enhance road accessibility in some Kisumu communities (Khanani et al., 2021). During the last County Integrated Development Plan of Kisumu period (2018 – 2022), 15 km of bitumen standard roads were constructed, significantly improved access and connectivity within informal settlements (CGK, 2022). However, overall, well-planned streets are still concentrated in the city centre.

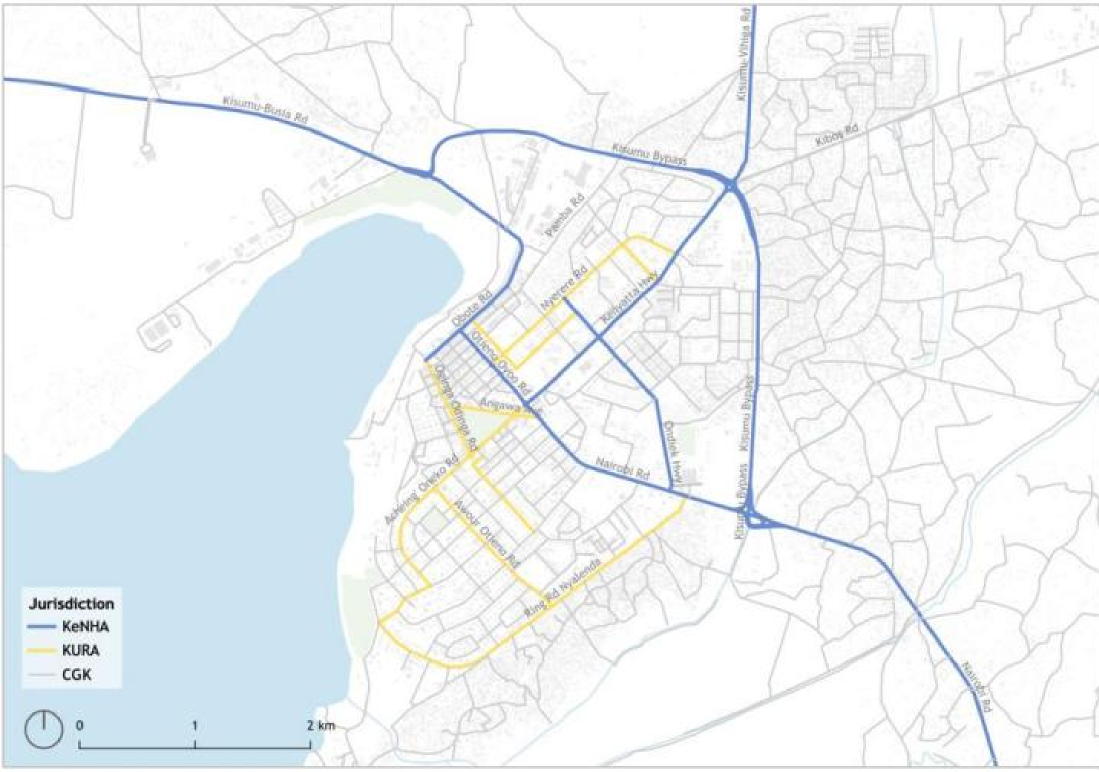


Figure 3.4 Street jurisdiction in Kisumu (County of Kisumu, 2020).

3.1.3 Urbanisation and infrastructure development in Kigali

3.1.3.1 Urban layout

Like Kisumu, Kigali also originated from trading posts established during the colonial period (Manirakiza et al., 2019). Along with the city's expansion, in 1964, the first Conceptual Master Plan for urban planning was presented. Later, in 1990, the Préfecture de la Ville de Kigali (PVK) was introduced to manage Kigali, which had expanded from 70 to 112 square kilometres. By the time the Mairie de la Ville de Kigali replaced the PVK in 2000, the area of Kigali had reached 314 square kilometres. Finally, in 2005, the name 'City of Kigali' (CoK) was officially given to this city, which had grown to 730 square kilometres and consisted of 3 districts, 35 sectors, 161 cells and 1,061 villages (Baffoe et al., 2020b, City of Kigali, 2020b, Manirakiza et al., 2019).

The configuration of urban space and population in Kigali was shaped mainly during the 1990s, in the post-genocide period (Esmail and Corburn, 2020). In the two decades from 1991 to 2022, Kigali's population grew from 1.3 million to about 1.75 million (NISR, 2023a, ONAPO, 1991). According to the 2022 Rwanda Population Census, most of them are located in the Gasabo district (50.4%) (NISR, 2023b). The high urban population proportion of 86.9% also makes the City of Kigali the most urbanised province in Rwanda (NISR, 2023b). Meanwhile, Kigali experienced a rapid increase in the built-up areas between 1984 and 2016, with a net change of 887.9% in high-density buildup areas (Mugiraneza et al., 2019), with the residential area of Kigali increased by 20.6% from 2013 to 2018 (City of Kigali, 2020d). As with SSA urbanisation described in the literature review (section 2.2.1), in Kigali, rapid urbanisation is also accompanied by informal area expansion (known as *akajagari* in the local language) or unplanned settlements as Kigali officials prefer to call them (NISR, 2018). As for the population currently living in Kigali, about 77.3% of the households live in informal settlements (NISR, 2018), about 79% of the total population (Hitayezu et al., 2018). It is widely recognised that the emergence of informal settlement in Kigali is linked to the influx of post-genocide migrants. In the 2022 census, the total number of immigrants in Kigali reached 354,970 (NISR, 2023b). Hitayezu et al. (2018) estimated that 65% of adult residents living in unplanned settlements are immigrants who moved to Kigali for geopolitical or economic reasons. Due to the importance immigrants place on ease of work, these communities tend to be close to well-developed roads (Hitayezu et al., 2018, Uwizeye et al., 2022). Meanwhile, they are excluded from the small slope areas that are less hazardous because of Kigali's hilly terrain (Baffoe et al., 2020b, Nduwayezu et al., 2021, University of Rwanda, 2018). This pattern is consistent with the common distribution characteristics of urban slums in the global south as summarised by Kuffer et al. (2017). Overall, the city follows a concentric urban land-use model, extending from the central business district (CBD) areas to informal settlements on the outskirts, with some high-end housing scattered on the outermost periphery (Nduwayezu et al., 2016, University of Rwanda, 2018). Baffoe et al. (2020) further divided Kiagli's neighbourhoods into three typologies: planned, unplanned /informal, and a mixture of the two (neighbourhoods resulting from upgrading, degradation or amalgamation). The distribution of informal settlements is illustrated in Figure 3.5.

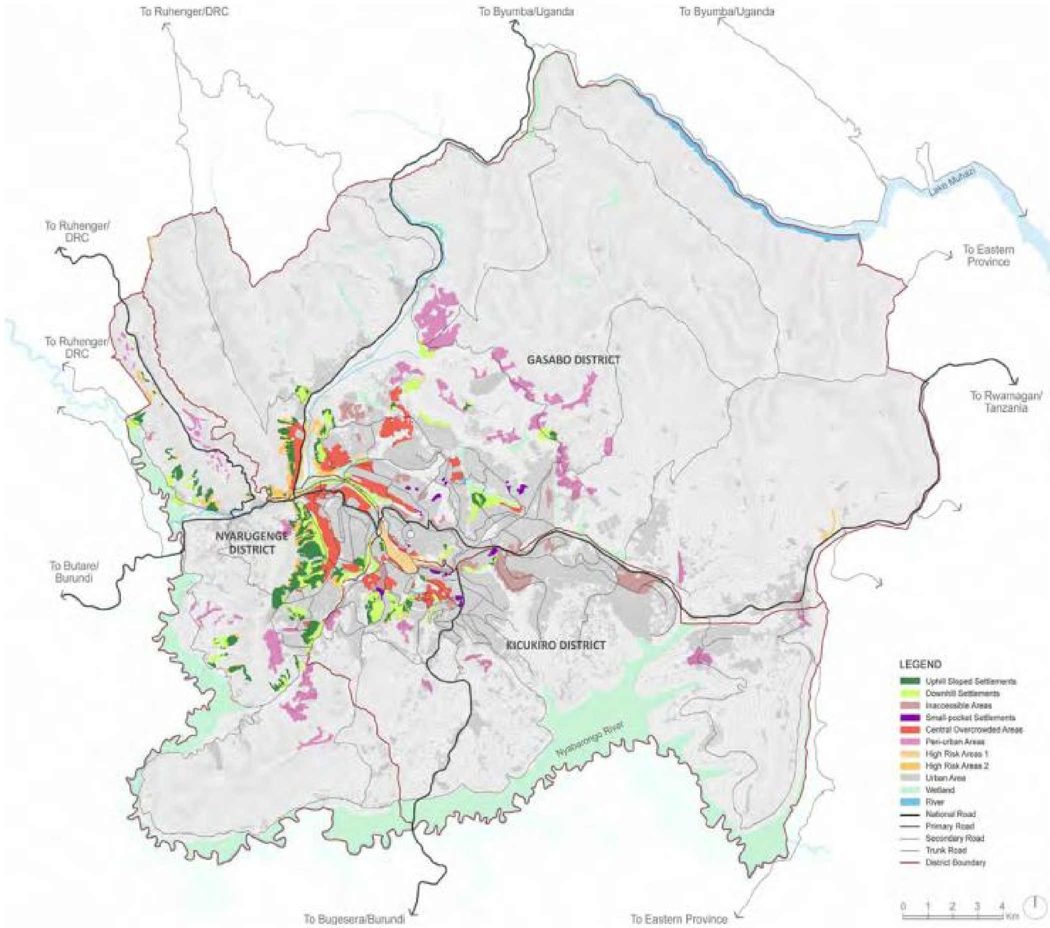


Figure 3.5 Map of informal settlements in Kigali (City of Kigali, 2020a). The UN-Habitat and the City of Kigali have classified these settlements into seven types.

The dominant categories include the uphill sloped settlements and overcrowded settlements located in Nyarugenge, which is situated in the city centre.

The distribution of informal settlements in Kigali, while following the general pattern of urban expansion in SSA, was also influenced by colonial-era policies. Esmail and Corburn (2020) note that some informal settlements can date back to the colonial period communities and that they coexisted with planned areas until the Rwandan government intervened. Based on the National Informal Settlements Upgrading Strategy's classification of informal settlements (MININFRA, 2015a), the informal settlements with long histories include Akabahizi, Munanira, Kimisagara, Gatenga, Karambo, Nyanza. Nyagatovu, Nyabisindu, Kangondo I, and Kangondo II (Uwizeye et al., 2022). In contrast, areas that were the focus of the French Ministry of Cooperation (e.g., Kicukiro, Remera, Kimihurura, and Kacyiru) are now the home of the high-income groups (Benken, 2017).

3.1.3.2 Water services and roads

Before 2014, Kigali's water provision was the responsibility of EWSA, the Energy, Water and Sanitation Authority (University of Rwanda, 2018), and now the Water and Sanitation Corporation (WASAC) is in charge of urban water supply (City of Kigali, 2020a). The main water sources of the Kigali water supply system are the Yanze River, Nyabarongo River and Lake Mugesera. Water is distributed to consumers from these sources through WASAC's 494km of pipelines, which cover most of the built-up areas in the city. Urban areas close to the city centre, such as Muhima, Nyarugenge, Gitega, Kimisagara, Kimihurura, Gikondo, etc., have the most extensive piped water network. In the past, the water network extension was not planned according to the future land use or growth. Hence, many existing pipelines do not follow the road alignments and encroach into property boundaries. This has made maintenance work challenging to carry out. At the same time, the undulating terrain also adds cost to any expansion (City of Kigali, 2020a).

Although risks also exist in Kigali's basic service systems, such as water and sanitation (Tsinda et al., 2020), the overall performance of water services in Kigali remains impressive. In 2018, WASAC reported that 86% of the Kigali population had access to water within 200m of travel, while the proportion of the population able to access household water connection reached 50% (NISR, 2018). The Rwanda DHS report shows that in the City of Kigali, the population with access to improved drinking water sources (piped water, protected well/spring, etc.) reached 97.4% (NISR, 2021). Hitayezu et al. (2018) estimate that 87% of households in unplanned settlements have access to improved water sources, with the value affected by the distance from the main road. The well-developed water system in Kigali can be explained by the strategy of its master plan, which is that the expansion of water and other infrastructures should always align with the city's overall development plan. Uwizeye et al. (2022), Benken (2017), University of Rwanda (UR) (2018) and Hudani (2020) all find that Kigali has deployed urban plans that incorporate a range of policies (e.g., the National Urbanisation Policy (MININFRA, 2015b), and the Vision 2020 (GoR, 2012)), which set strict standards for urban environment in order to form "the Singapore of Africa". The government expects that the transformed Kigali can serve as an intrinsic economic driver and hopes that inclusive policies can unite the city's inhabitants during the post-genocide period (Manirakiza et al., 2019). We can see this in Bolin (2019)'s description of the changes in Kigali, which Esmail and Corburn (2020) summarise as "a large technological system, with its harmonised edifices of zoning and legal provisions, which segment and integrate space and social life." The implementation of these policies has objectively enabled the residents of Kigali's informal settlements to access public services such as water, sanitation, electricity, and land registration. Of course, according to the master plan of Kigali, providing basic services to the residents of the informal settlements does not justify the presence of the informal settlement in the "modern city with high standards" (Esmail and Corburn, 2020, Uwayezu and de Vries, 2020). At the same time, paid services, such as healthcare and education, remain unaffordable for informal settlement residents (Hitayezu et al., 2018, Uwizeye et al., 2022). Thus, while the small drinking-water supplies (SDWS) projects similar to the DMM also exist in Kigali (Herschan et al., 2023, University of Rwanda, 2018) , Kigali prefers the formal way to cope with the increase in demand due to rapid urbanisation (City of Kigali, 2020a).

For the same reason, Kigali's road network (Figure 3.6) exhibits the best connectivity among all roads in Rwanda. The construction and management of these roads and their traffic fall under the jurisdiction of the Ministry of Infrastructure (MINIFRA) and its subsidiary agency, the Rwanda Transport Development Agency (RTDA). They are focusing mainly on improving the standards of the road network, such as paving roads and establishing clearer road classification and design criteria (City of Kigali, 2020a).

The layout of existing roads is recognised as being primarily influenced by topography and strongly linked to the city's layout (Dufitimana and Niyonzima, 2023, Hitayezu et al., 2018). It can be stated that the expansion pattern of Kigali City is horizontal, and the expansion and infrastructure development largely adhere to a similar pattern. These newly expanded areas tend to border existing urban core areas. Correspondingly, streets of new areas evolve along the city's main roads (Nduwayezu et al., 2021).

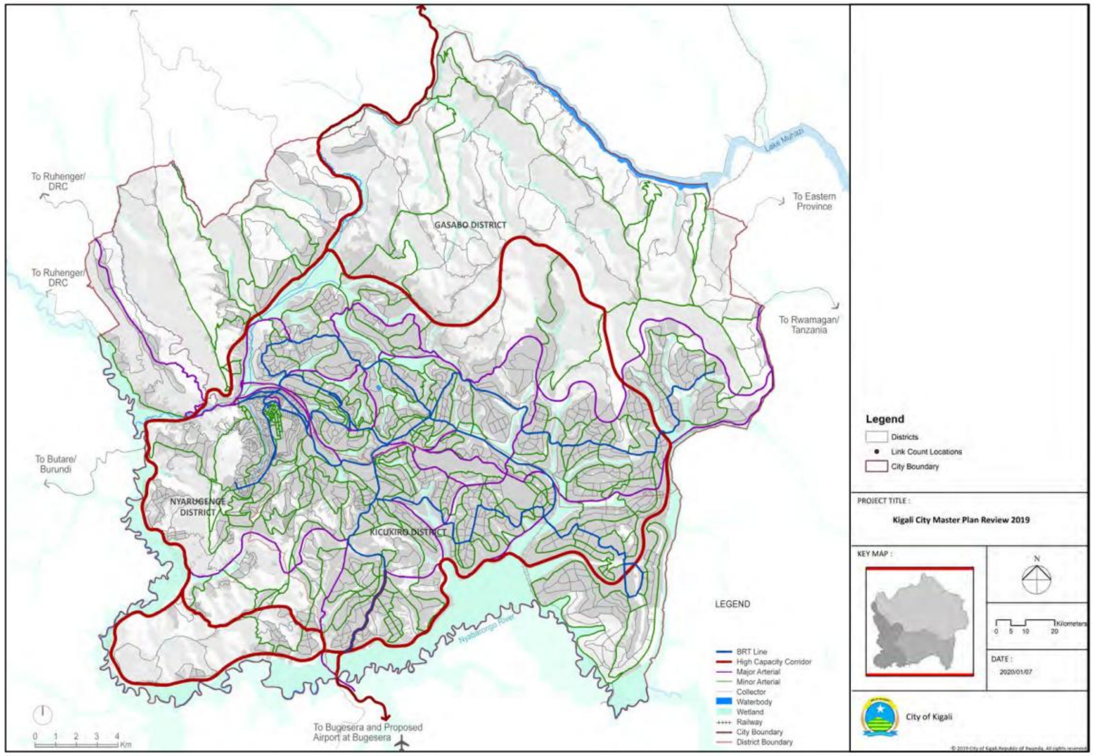


Figure 3.6 Kigali road map (City of Kigali, 2020c). The roads are classified based on capacity and function. Under this classification, the roads in Kigali are primarily encompassed by the Kigali Ring Road, which is classified as High Capacity.

3.2 Software and packages used

To support the various stages of analysis in this research, a range of software and packages were employed.

ArcGIS Pro 2.7.3 (Esri) was used extensively for geospatial data editing and analysis. This included the identification of topological errors in the pipeline network using the topology validation tool (Chapter 4), spatial statistical analyses such as Global and Local Moran's I (Chapter 6), and interpolation methods including inverse distance weighting (IDW) and kernel density estimation (KDE) (Chapter 6). The calculation of population-weighted centroids (Chapter 6), the co-location analysis of pipelines (Chapter 4), and the estimation of 2SFCA accessibility (Chapter 6) were also carried out within this platform. All maps presented in the analytical chapters were generated using ArcGIS Pro except for Figure 4.5.

AccessMod 5 (Ray and Ebener, 2008) was employed to simulate travel paths as part of the accessibility analysis in Chapter 6.

For network-based analysis, a combination of R and Python packages was used. The R package *sf* (Pebesma and Bivand, 2023) and *sfnetworks* (van der Meer et al., 2021) were applied for graph construction and smoothing (Chapter 4, Section 4.2). Package *networkx* (Hagberg et al., 2008) and *igraph* (Csardi and Nepusz, 2006) libraries were used for converting between primal and dual graph representations and for calculating graph metrics (Chapters 4 and 5). The *shp2graph* package (Lu et al., 2018) was employed to detect and visualise self-connected components within the network (Chapter 4). The *OSMnx* package (Boeing, 2017) facilitated the downloading and preparation of OpenStreetMap road data for graph-based analysis (Chapter 4). The Infomap algorithm (Rosvall et al., 2009) was applied via the MapEquation platform (https://www.mapequation.org/infomap/) (Chapter 5). Finally, the *aricode* package (Julien Chiquet, 2024) was used to compute partition similarity indices (Chapter 5).

Two statistical packages also supported specific analytical tasks. The *blandr* package (Datta, 2024) was used to conduct Bland–Altman analyses in Chapter 6, while *fitdistrplus* (Delignette-Muller and Dutang, 2015) was used for fitting statistical distributions in Chapter 5.

Chapter 4 Geometric and topological convergence of urban networks

The performance of water pipeline networks is closely related to their morphology. In addition to affecting network robustness (Agathokleous et al., 2017, Bentes et al., 2011, Wang et al., 2019), the topology of a water pipeline network also influences water transmission efficiency (Torres et al., 2017, Yu et al., 2024) as well as the complexity of management and maintenance (Godfrey and Howard, 2005a, Lagerberg, 2016). As highlighted in Section 2.3.1 of the literature review, urban infrastructure networks are a product of the industrial-era urbanization process in Europe (Coutard, 2008, Coutard and Rutherford, 2015), designed primarily to ensure equal access to services for all urban residents. Consequently, infrastructure network expansion follows two key processes: area expansion and densification (Yang et al., 2017), which involve connecting new settlements and enhancing coverage in already serviced areas.

However, this concept faces challenges in Africa. Cities in Sub-Saharan Africa (SSA) experience a distinct form of informal urbanization (see Sections 2.1 and 2.2.1 of the literature review). This not only differentiates their economic growth patterns from historical European urbanization but also alters spatial development, which in turn affects infrastructure network performance (Batty, 2012). A study on SSA urban morphology using Accra, Ghana, as a case study found that urban growth in SSA is characterized by increasing complexity and fragmented spatial development (Korah et al., 2019). Cobbinah and Niminga-Beka (2017) and Agyemang et al. (2019) noted that this process is often accompanied by unplanned land-use changes, reflecting spontaneous, inefficient, and poorly regulated development. These characteristics further affect infrastructure network morphology and performance. A striking example is the water network in SSA slums. The emergence of slums is closely linked to urbanization outpacing government capacity, resulting in highly informal settlements characterized by self-organized management and a lack of formal planning (Azunre et al., 2022, Okyere et al., 2017, Sheppard et al., 2020). While definitions of slums vary, descriptions across many developing countries consistently highlight their dense population and built-up environment (Criqui, 2018, MLHUD, 2008, Moreno, 2003, Nolan, 2015, UN-Habitat, 2006). These characteristics are closely associated with urban water supply challenges (Azunre et al., 2022, Nyika and Dinka, 2023, Richmond et al., 2018, Sinharoy et al., 2019). As discussed in Section 2.3.1 of the literature review, three key challenges inhibit slum water pipeline network development:

1. The unplanned buildings and road networks in slums limit the space available for pipeline laying, thereby making the pipelines have inefficient topology.

2. Insecure land ownership, low-income levels, and governance failures further weaken pipeline planning and maintenance in slums.
3. Due to colonial history, the infrastructure baseline in slums differs significantly from that in formal areas, increasing the cost of further pipeline expansion.

Therefore, assessing water services in SSA cities inevitably requires considering the unique characteristics of their urban morphology.

Urban morphology is often evaluated through the geometry and topology of street networks. In urban morphology studies, cities are understood as spatial organisations of different elements—primarily plots, buildings, and streets—forming recognisable patterns that influence urban life (Araújo de Oliveira, 2022b, Scheer, 2015). Among these elements, streets are particularly emphasised. Streets serve as the structural framework of urban organisation, integrating other elements. Compared to street blocks, plots, and buildings, streets exhibit greater stability throughout urban evolution while also serving as a key representation of a city's structural characteristics (Araújo de Oliveira, 2022a, Wang and Gu, 2023). Streets connect new settlements while linking private spaces, simultaneously delineating public spaces that facilitate the flow of materials and energy within the city (Kropf, 2014). Therefore, urban amenities and infrastructure networks often follow similar spatial patterns and evolutionary trajectories to street networks. Consequently, street networks are widely regarded as simplified schematic representations of urban contexts and are frequently used to classify and analyse different urban morphologies (Cardillo et al., 2006, Zhang et al., 2023). Numerous studies have employed street network analysis to investigate various aspects of urban systems. For instance, Dingil et al. (2018) analysed road data to examine the relationship between infrastructure accessibility and socioeconomic indicators across different urban typologies. Serra et al. (2016) described six decades of morphological changes in the Oporto Metropolitan Area through an analysis of street network geometry and topology. Buhl et al. (2006) identified self-organised urban settlement patterns through street network analysis. Negadi et al. (2023) assessed urban connectivity by examining fractality and connectivity of street networks. Additionally, Jang et al. (2024) explored the correlation between street characteristics and urban vibrancy, with population mobility serving as the indicator of vibrancy in this study. Their findings demonstrate that street features influence the vibrancy of different age groups at various times of the day. This human mobility pattern is also considered closely linked to urban morphology (Kang et al., 2012). Some studies have further extended this approach to examine the impact of slums on urban morphology (Dovey et al., 2020, Kolowa et al., 2024).

The spatial distribution of roads has also been recognised as being closely related to the structure and performance of pipeline networks. For example, Abdel-Mottaleb and Zhang (2020) assert that there is an unquantified dependency between urban water supply and transportation networks. Debón et al. (2010) and

Așchilean et al. (2018) highlight that pipeline failure risks are associated with road traffic loads. Mair et al. (2017) further analyse the shared characteristics of street networks and urban water infrastructure networks. However, their study is based on three anonymised cities and does not explore the role of urban morphology in shaping these networks. Additionally, from the network science perspective, urban infrastructure systems—including water, electricity, and transport networks—can be analysed as an interdependent multilayered model (Munikoti et al., 2021). Despite this, no existing studies have explicitly established a correlation between urban morphology and the structural patterns of both road and water pipeline networks. Lorenz et al. (2021) attempted to compare pipeline network attributes across cities, yet their analysis was not based on a detailed examination of urban morphology. Moreover, their analysis did not examine pipeline attributes beyond resilience. Another attempt to analyse the relationship between road and pipeline networks in different cities comes from Abdel-Mottaleb and Zhang (2020). However, their study relied solely on synthetic pipeline networks, focusing on the properties of the interface network formed by roads and water networks. Another relevant study is that of Chegini and Li (2022), which explored the topological relationship between street networks and belowground urban stormwater systems. However, the stormwater infrastructure considered in their work is specifically designed to collect runoff and protect streets from flooding, and is thus inherently more tightly integrated with the street layout. As a result, the observed correspondence in centrality and spatial positioning between streets and pipelines in their analysis is context-dependent and of limited generalisability. More importantly, no study has specifically investigated the relationship between the unique morphological characteristics of SSA cities and their water distribution networks.

The performance of a pipeline network can be measured from multiple perspectives, including reliability, risk, vulnerability, and resilience (Shuang et al., 2019, Soldi et al., 2015). However, the calculation of these metrics requires different techniques and an understanding of the hydraulic characteristics and mechanics of the pipeline (Jensen and Jerez, 2018, Yazdani and Jeffrey, 2011). Such an approach often faces many challenges:

1. The pipelines are located underground, and it is difficult to determine the condition of the network visually, thus requiring more effort to obtain data (Yu et al., 2024).
2. The pipeline network is a complex system, and its assessment is often accompanied by a number of constraints, such as dependencies between components, water pressure requirements, and urban planning requirements, making the analysis difficult (Jensen and Jerez, 2018).
3. As the size of the network increases, these constraints become increasingly complex, and new variables are introduced (Herrera et al., 2015).

Therefore, measuring large urban water and sanitation networks requires a large amount of data as model inputs (Perelman and Ostfeld, 2011, Torres et al., 2017). The situation is even more complex in low- and middle-income countries like most of those in SSA. Consequently, assessing the pipeline configurations in SSA slums using this method presents significant challenges.

This is where the advantages of graph theory come into play, as it captures the structure of pipelines associated with the performance of water services. The coverage of water network directly affects residents' proximity to urban water and sanitation services, while the network structure influences efficiency and robustness of the water supply. Therefore, compared to traditional methods, graph assessments are topological in nature and can yield valuable insights into network performance. Given the spatial and policy differences between cities and between slum and non-slum areas, if there are differences in their pipeline networks, graph metrics can provide performance indicators relevant to the Sustainable Development Goal (SDG) 6 and water policies. The results of this assessment are topological, and therefore the results obtained are comparable between different areas.

Therefore, this chapter examines the relationship between the topology of pipeline networks and SSA urban morphology from a global perspective, analysing their correlation from both geometric and topological perspectives. It aims to address two major research gaps:

1. Urban morphology studies have predominantly focused on cities in the Global North. As a result, variations in urban form across cities in the Global South, particularly SSA, have received limited attention.
2. Road networks serve as the structural backbone of cities, influencing other infrastructure networks. However, few studies have explicitly linked the morphological characteristics of road networks with pipeline networks. Furthermore, existing research has not considered the impact of slum settlements on this relationship.

To bridge these gaps, this chapter employs a co-location approach to analyse the morphological differences between roads and pipelines occupying the same spatial locations. Given that road networks serve as indicators of urban morphology, this approach aims to identify the relationship between pipeline networks and urban form, specifically by:

1. Quantifying the differences in road networks—and by extension, urban morphology—across cities with varying types of slum settlements.
2. Assessing the association between urban morphology and pipeline networks and potential implications for water service provision.
3. Comparing the commonalities between road and pipeline networks across case study SSA cities.

4.1 Overview of methodology

We here examine the commonalities of three infrastructure networks in Kisumu—road, water pipeline, and sewer pipeline networks—as well as the road and water pipeline networks in Kigali, via analysis of their topological and geometrical similarities (Figure 4.1). The urban infrastructure networks within the study area are first converted into dual graphs and compared using graph theory approaches. In addition to employing graph theory metrics to analyse general topological characteristics, the dissimilarity method proposed by (Schieber et al., 2017) was applied. Subsequently, centrality metrics (degree, betweenness, closeness) obtained from graph analysis were mapped back into geometric space. Through co-location analysis, overlapping features of water and sewer pipeline networks and road networks, as well as the correlation between their centrality metrics, were identified.

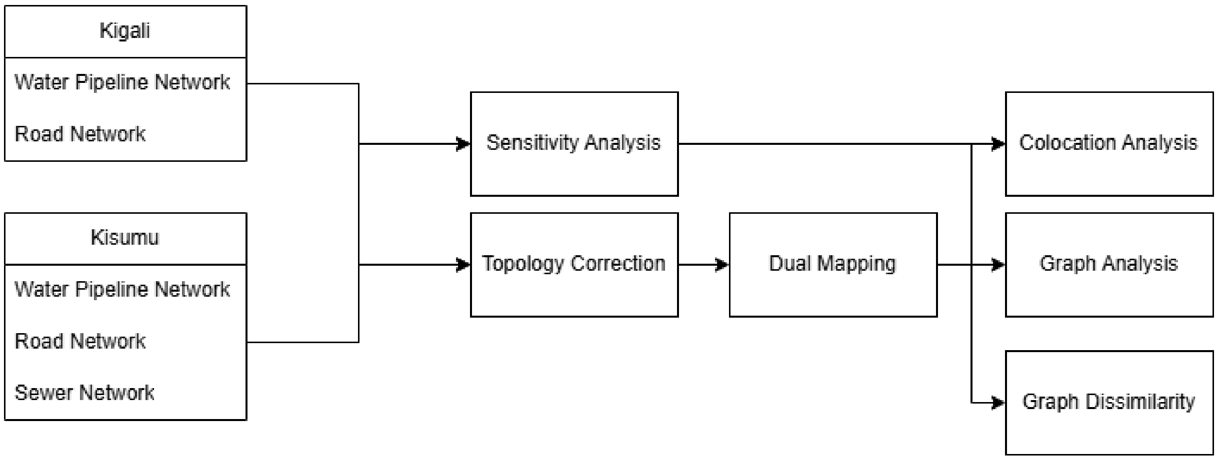


Figure 4.1 Workflow for comparing topological characteristics of road, water and sewerage networks in Kisumu and Kigali.

4.2 Data preprocessing

Due to data availability constraints, sewer network data for Kigali was not accessible. Therefore, this study utilized water, sewer, and road networks from Kisumu, along with water and road networks from Kigali. The road networks primarily served as a reference for identifying urban morphology. To enable the application of graph methods, which are widely used in network morphology studies, network data were preprocessed before conversion to graph format. This preprocessing included network feature identification, topological correction, graph smoothing, and dual mapping.

4.2.1 Downloading and preprocessing road network data

In the study area, multiple data sources are available for road networks, providing more options than for pipeline networks. As of February 4, 2024, these sources include the Kenya Roads Board (KRB) (KRB, 2021), Rwanda Transport Development Agency (RTDA) (RTDA, 2022), Global Roads Inventory Project (GRIP) (Meijer et al., 2018), and OpenStreetMap (OSM) (OpenStreetMap contributors, 2023). KRB and RTDA are national road management authorities in Kenya and Rwanda, respectively. The road data from KRB, collected through a national survey in 2018, covers only the main roads in Kisumu, with a total length of 519,524 meters. Similarly, RTDA's dataset includes only National Roads, Class 1 District Roads, and roads in the City of Kigali and other urban areas. The GRIP Road database, covering the study area, integrates multiple data sources, including OpenStreetMap. However, due to its primary focus on environmental and biodiversity research such as GLOBIO (Schipper et al., 2020), the level of detail in GRIP's road data is insufficient for other research applications. In Kisumu, GRIP's road length is 415,016 meters, while in Kigali, it spans 2,024,631 meters. Comparatively, the OSM road network in Kisumu has a length of 1,681,830 meters, and in Kigali, the OSM road length is 5,022,458 meters, including a greater proportion of local roads. Since OSM's total road lengths are greater than those of other data sources, and the spatial distribution is more uniform, this chapter uses OSM as the road data source. The Python package *OSMnx* (Boeing, 2017) was employed to retrieve OSM road maps within the boundaries of Kisumu and Kigali as of February 4, 2024.

The *OSMnx* package provides built-in functionality for processing downloaded road networks from OSM into planar graphs and performing topological corrections. Since OSM data includes 3D information, such as bridges, *OSMnx*:

1. processes the network into a planar graph, preserving only the 2D projection of the road network while excluding 3D intersections—a standard approach in road network analysis (Boeing, 2018).
2. converts the 2D network into a graph format supported by the *NetworkX* package (Hagberg et al., 2008). It represents road network as an undirected primal graph, where road junctions correspond to graph nodes, and road segments are represented as edges(Añez et al., 1996).
3. In this study, street angles and road attributes were not considered in the modelling process to ensure consistency with the pipeline network representation.

4.2.2 Preprocessing of water pipeline networks

Kisumu's water pipeline network data was obtained from KIWASCO in September 2020, following a formal data request (Figure 4.2). The raw dataset comprised 21,448 polyline features depicting pipelines with a total length of 538,559 meters. In contrast, the sewer network was comparatively smaller, comprising 253 polylines with a total length of 90,467 meters (Figure 4.3).

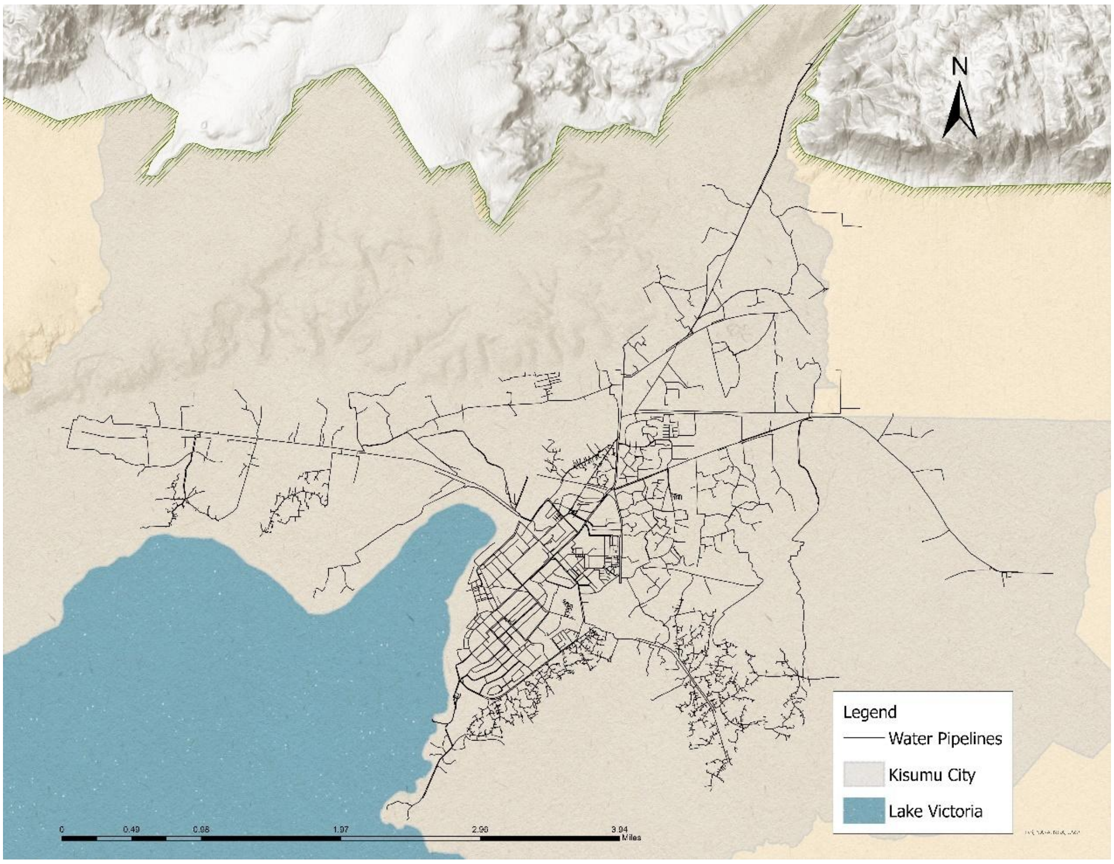


Figure 4.2 Piped drinking-water network in Kisumu City. Boundary data created based on the boundary map produced by the American Red Cross (American Red Cross, 2019)).

Figure 4.3 Sewer network in Kisumu City. Boundary data created based on the boundary map produced by the American Red Cross (American Red Cross, 2019)).

Kigali's water pipeline network data was obtained from the Water and Sanitation Corporation (WASAC) via an open data repository, WaterGIS (https://github.com/watergis). The data, updated in 2020, comprises a larger network than that of Kisumu, including 11,072 line segments with a total length of 1,728,445 meters (Figure 4.4).

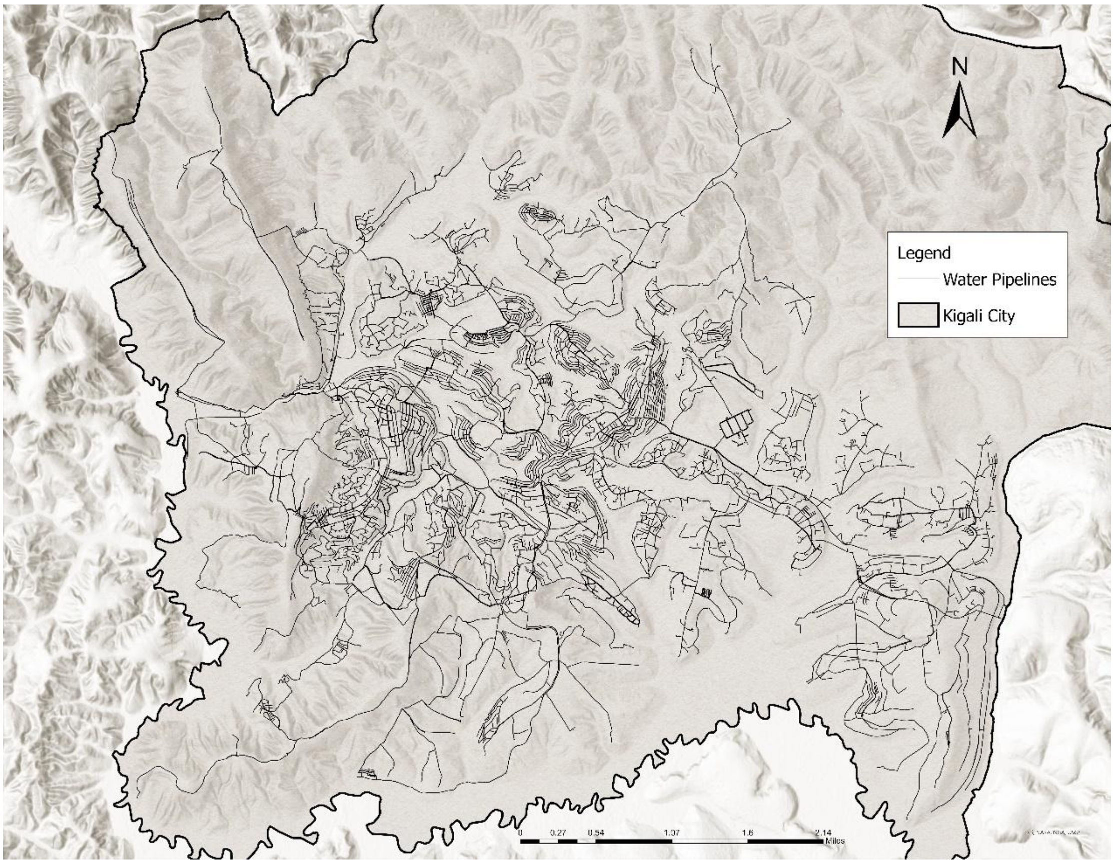


Figure 4.4 Piped drinking-water network of Kigali City. Boundary data extracted from the administrative level 1 boundaries of Rwanda in the Common Operational Datasets (CODs) (OCHA, 2023).

4.2.2.1 Manual topology correction for pipeline data

Piped network layers were provided in non-topological shapefile format, so topology-building was necessary. Theoretically, a city's pipeline network should have no pipelines that are disconnected from the rest of the network. Therefore, it is assumed that the study area contains only a single connected network; any isolated pipelines present are considered digitising errors. R's *nt.connect* function in its *shp2graph* library was used initially to identify self-connected parts within each network (Lu et al., 2018).

As shown in Figure 4.5, there are many self-connected parts in Kisumu's water pipeline network with similar problems in Kigali. The reasons include topological problems such as overshoots, overlapping pipelines, misalignment of endpoints and interruptions in line segments (see examples in Figure 4.6). Since the modelling of pipeline networks follows principles outlined in section 2.4.1, graph metrics will be misleading when modelling with disconnected network components. Kisumu and Kigali's piped networks both contain topologically incorrect points and lines, and the node and edge lists cannot be used to build a graph without prior correction.

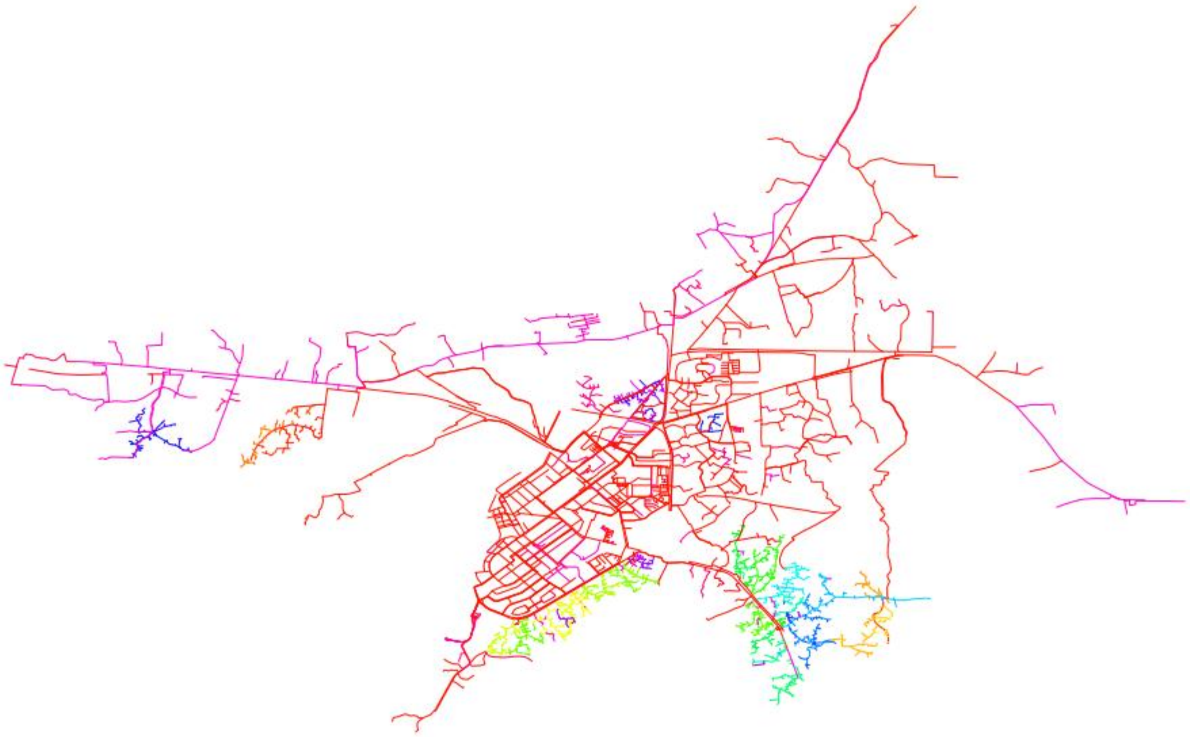


Figure 4.5 Self-connected parts in the Kisumu network before topological correction, with 1,134 parts plotted in different colours.

However, on one hand, *OSMnx* cannot be used to correct topology errors in pipelines as it does for OSM roads due to incompatibility with the data structure required by the *OSMnx* package. On the other hand, ArcGIS Pro's automated topology correction workflows have limitations when applied to urban networks. ArcGIS Pro software automatically identifies and handles polyline topological errors including overlaps, intersects, and dangles. Take dangles, for example, as shown in Figure 4.6 (A), identified by ArcGIS where endpoints of lines do not touch other line segments, both real-world consumer endpoints and topologically incorrect breakpoints. The latter are abundant in Kisumu (Figure 4.6 (B)). However, Figure 4.6 (A) also shows that in residential areas, the distance between consumer endpoints is often small as well, so these pipelines without topology errors will also be identified as dangles and be trimmed or extended in ArcGIS. Since this automated topology correction creates errors, all pipelines were therefore corrected manually.

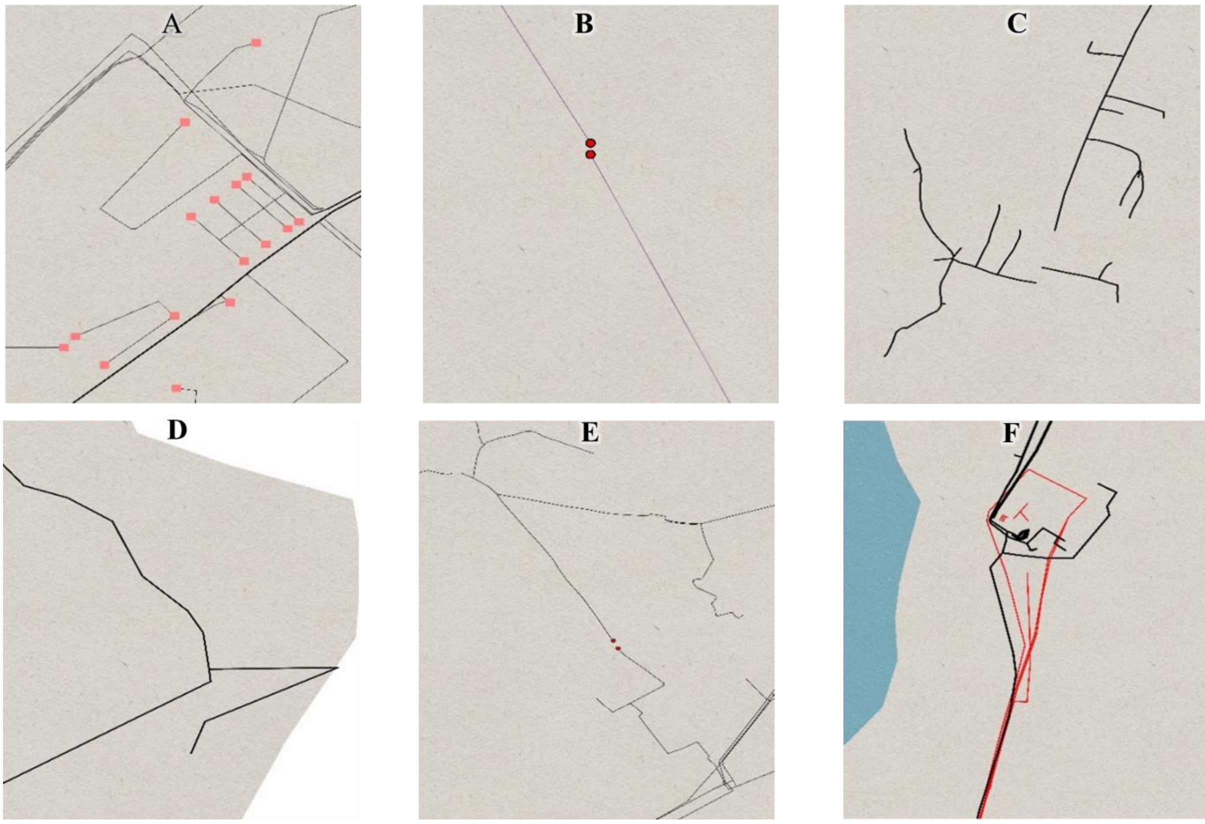


Figure 4.6 Common topological errors in water pipeline networks, illustrated with examples from the Kisumu pipeline network. A. dangles; B. breakpoints in pipelines; C. isolated pipelines; D. overshoots with large angles; E. isolated parts with small gaps from the main; F. closed network segments.

After diagnosing potential topological errors in the networks, these errors were subsequently corrected via the strategies shown in Table 4.1. The principle of the correction is to make as few changes as possible to the original data. The city's pipeline network shows a certain pattern: the main pipelines distributed along the main roads extend in parallel and are rarely connected to each other. At the same time, the distribution of pipelines in residential areas is often quite regular. These are also the conditions to be considered when making amendments.

Table 4.1 Potential Topology Errors and Corresponding Correction Strategies in the Study Networks

Potential Topology Errors	Examples	Correction Strategy
Breakpoints	Figure 4.6 (B)	To identify these, a tolerance distance of two meters is applied; gaps smaller than this threshold are considered topological errors and are resolved by reconnecting the segments.
Isolated pipelines/self-contained segments	Figure 4.6 (C)	
Overshoots with large angles of extension	Figure 4.6 (D)	Short overshoots or undershoots that deviate significantly from the angle extending an existing line feature are considered to be deliberate on the part of the cartographer and, therefore, very unlikely to be topological errors and not corrected.
Isolated parts with small gaps from the main networks	Figure 4.6 (E)	If an isolated pipe can be connected to the nearest network section by extending the undershoots, the connection is made, even if the distance is considerable (e.g., 10m).
Closed network segments	Figure 4.6 (F)	A pipeline running parallel to the nearest pipe that cannot be connected is assumed to belong to a different system and is temporarily removed.

4.2.2.2 Transforming pipeline data to graph format

During the preprocessing steps, the pipeline networks were first transformed into an undirected graph through primal mapping (Figure 4.1). Although the direction of water flow is critical for hydraulic simulations of the network, detailed information (e.g. on pressure head, pipe diameter and gradient) required for hydraulic analysis is often lacking. As a result, pipeline networks are frequently treated as undirected graphs, a format in which hydraulic direction is not considered (Boccaletti et al., 2006, Hwang and Lansey, 2017, Meijer et al., 2018, Yazdani and Jeffrey, 2012a, Yazdani and Jeffrey, 2012b). Primal mapping is an intuitive modelling method and is commonly used for graph analyses of water distribution networks (Yazdani and Jeffrey, 2011). In primal mapping, specific locations (e.g. pipeline and road intersections, reservoirs, consumers and pumps) are represented as nodes and pipelines are defined as edges between nodes within a network (Yazdani and Jeffrey, 2011, Yazdani et al., 2011).

To convert the network into a graph, it is necessary to determine which elements in the raw network should be retained based on factors such as their attributes and the angles between them (Marshall et al., 2018, Stavroulaki et al., 2017). Treatment varies from modelling methods such as space syntax (Hillier and Hanson, 1989), the street name approach (SN), intersection continuity negotiation (ICN) (Porta et al., 2006), and Hierarchical Intersection Continuity Negotiation (HICN) (Masucci et al., 2014). Nevertheless, studies of road and pipeline networks have not considered whether edge attributes or angles between features are the most appropriate basis for graph modelling (Giudicianni et al., 2018, Hwang and Lansey, 2017, Marshall et al., 2018, Pagano et al., 2019, Yazdani and Jeffrey, 2012a). Since the case study networks provide little information on components besides pipelines, only pipelines and intersections were considered in modelling:

1. Endpoints and intersections of line features were identified
2. And pipelines with the same ID were merged and then split at the intersections.
3. The resulting shapefile was then read as an *sf* class object, storing the node and edge information using R's *sf* library (Pebesma and Bivand, 2023).

4.2.3 Graph smoothing

Prior to analysis, graph smoothing (sometimes called pseudo-node removal) was undertaken. The raw graph object contains vertices that are likely to contribute little to graph structure and hinder the processing and interpretation of the graph (Ersoy et al., 2011, Hennessey et al., 2008, Hwang and Lansey, 2017, Ruan et al., 2011, Yu et al., 2024). *OSMnx* provides a built-in method for simplifying road networks, which eliminates nodes that do not correspond to intersections or endpoints while merging edges and preserving geometric attributes (Boeing, 2017). However, the simplification algorithm in *OSMnx* is specifically designed for road networks formatted according to OSM data structures, incorporating real-world connectivity constraints. To maintain consistency in network simplification across different infrastructure types, this study instead adopted the R package *sfnetworks* (van der Meer et al., 2021) for graph smoothing. The package implements a simplification method by retaining only the endpoints of line segments, the point from which an edge self-loops (which are excluded in the simplified graph; see Section 2.4.1), and the intersection of multiple edges where at least one of the edges continues through the intersection (Figure 4.7).

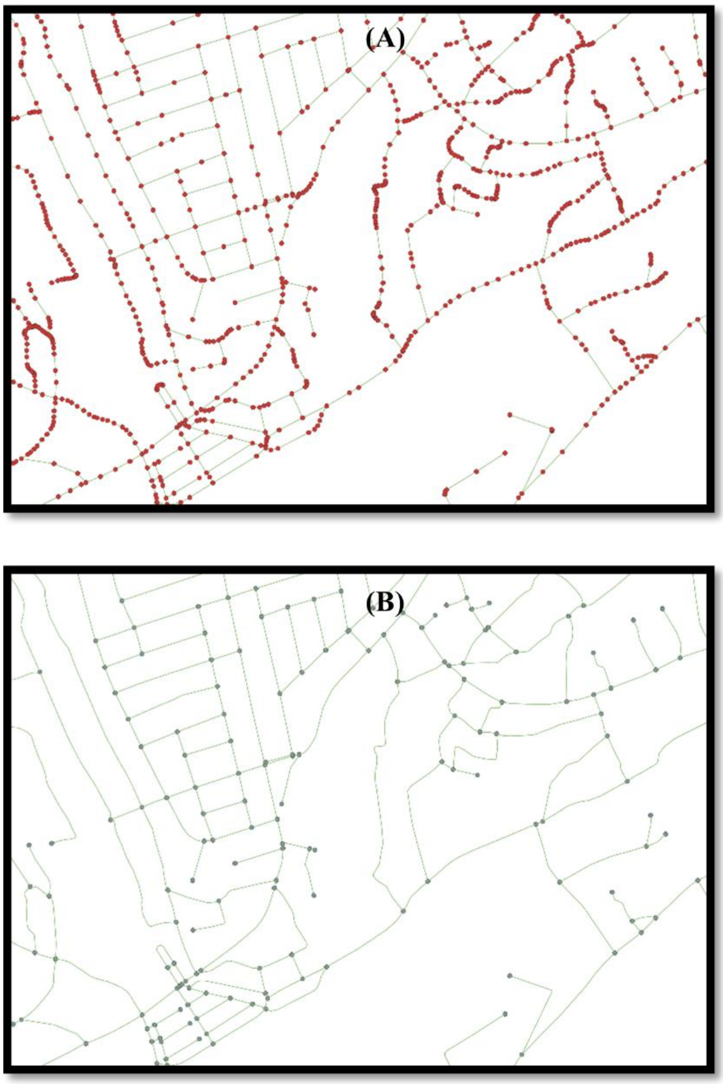


Figure 4.7 (A) Original street network obtained from OSM. (B) Simplified network after pseudo node removal.

4.2.4 Dual mapping of infrastructure networks

After obtaining the corrected pipeline and road networks, the primal graphs were converted into dual graphs using *igraph*. In dual mapping, the edges of the primal graph are represented as nodes, and connections are established between these nodes if the corresponding edges in the primal graph share an endpoint. Mathematically, a dual graph is the duality of the primal graph, which gives it its name; however, in practice, the dual graph is not always a precise dual of the primal graph (Marshall, 2016). Consequently, some studies use the term line graph to refer to this type of graph (e.g. Gharaee et al. (2021)). For consistency, this study adopts the term dual graph throughout the text.

A disadvantage of the dual mapping approach is that when edges are mapped as nodes in a graph, the geometric characteristics are discarded. However, Porta et al. (2006) also point out the advantages of dual mapping. In the case of street networks, the number of streets (edges) that can be connected to an intersection (node) is finite in primal mapping. However, in a dual-mapped street network, the number of intersections (edges) that can be connected to a street (node) is infinite. This structure is closer to the networks in other research fields, thus allowing for comparisons between networks. Both modelling practices (primal and dual) are commonly used in studies of infrastructure networks (Adraoui et al., 2024, Giustolisi et al., 2019, Wéber et al., 2020, Zhang et al., 2017, Zischg et al., 2019). In cases where the focus is on the representation of network edges, the dual approach is more appropriate. This is particularly relevant for co-location analysis, where the goal is to compare the spatial overlap of networks. In such cases, analysing edges rather than nodes is a more reasonable choice, as the comparison is between roads and pipelines rather than specific locations within the network (e.g., intersections).

4.3 Graph-based analysis methods

The processed dual graphs were analysed to uncover their commonalities. This involved using graph metrics to describe network properties, applying D-measure to quantify network dissimilarity, and identifying overlapping features to extract their centrality patterns.

4.3.1 Graph metrics

The following graph metrics, selected from those introduced in the literature review (Section 2.4.1), were used to assess various aspects of efficiency and stability in urban networks:

- Average shortest path length: the average distance along the shortest paths connecting two nodes, compared to all possible pairs of nodes in the network.
- Average node degree: the average value of the number of nodes connected to each node in the network.
- Critical breakdown ratio: the threshold that a network will fail when removing nodes whose degree exceeds this value.
- Algebraic connectivity: the second smallest eigenvalue of the normalized Laplacian matrix, reflecting the robustness of the network.
- Spectral gap: the difference between the first and second eigenvalues of the adjacency matrix.
- Clustering coefficient: the presence of triangular loops in the network.
- Meshedness coefficient: the presence of loops in the network, with a higher value signifying a more connected network.

Given that this study focuses on urban morphology rather than on the functional performance of the pipeline network, indicators emphasizing morphological characteristics were selected. Of these metrics, average shortest path length is related to the efficiency of the pipelines. According to Meng et al. (2018), the average path length is strongly correlated with the diameter and graph radius and, thus, more representative; average node degree, critical breakdown ratio, and algebraic connectivity are all related to the fault-tolerance of the pipeline, also known as robustness. This study did not use density of bridges/articulation points as an indicator of robustness as they are not representative in water distribution networks (Meng et al., 2018). In addition, two parameters characterise the shape of the network, from which information about the efficiency and robustness of the network can be obtained: cluster coefficient, and meshedness coefficient.

Specifically, three graph metrics were employed to identify local network characteristics in the co-location analysis:

Degree centrality: This metric represents the degree of each node in the dual graph, indicating the extent to which a pipeline or road is connected to others. In other words, it reflects the positional importance of an edge within the network. For instance, a node with a degree centrality of 1 signifies that it has only one connecting edge, meaning the corresponding road/pipeline is located at the network's terminus.

Betweenness centrality: This metric measures how frequently a node appears on the shortest paths between any two nodes in the network. A higher value indicates that the corresponding pipeline or road plays a crucial role in maintaining network connectivity. If such an element is removed, other pipelines or roads would require longer topological paths to remain connected.

Closeness centrality: This metric is computed based on the shortest path distance from a node to all other nodes, reflecting the network's flow capacity at that node. As an example from Wang et al. (2011b), in an air transportation network, cities with higher closeness centrality—indicating proximity to other cities in the graph—tend to exhibit better economic performance.

4.3.2 Network similarity

There is widespread interest in the research and application of network similarities across various disciplines, including social sciences, medicine, and biology (Barabási et al., 2011, Coşkun and Koyutürk, 2021, Tarapata and Kasprzyk, 2009, Taylor et al., 2015). Generally, this issue can be considered as the graph isomorphism problem, which aims to measure whether two graphs are topologically equivalent. Over the past few decades, numerous techniques have emerged for studying graph similarity, primarily focusing on proposing effective and computationally efficient metrics. Measures include Graph Edit Distance (Bunke and Allermann, 1983), Singular Value Sequence (Xu et al., 2019), and Graph Fourier Distances (Lagunas et al., 2018), as well as algorithms based on node scores, such as the SimRank algorithm (Jeh and Widom, 2002) and Kleinberg's algorithm (Blondel et al., 2004). Their common concept is to treat the distance between graphs as a measure of their similarity (Shimomura et al., 2021).

Among these methods, this study employed the D-measure, proposed by Schieber et al. (2017) , to quantify the similarity/dissimilarity of infrastructure networks within the study area. The D-measure is primarily based on node-to-node connection distances and comprises three components: the first term compares the networks' distance distributions, capturing global topological differences; the second term assesses how each element is connected throughout the network; and the last term analyses differences in the way this connectivity occurs, through the examination of Katz centrality. Thus, the dissimilarity $D(G, G')$ between graphs *G* andG' is defined as:

$$D(G, G') = w_1 \sqrt{\frac{\mathcal{J}(\mu_G, \mu_{G'})}{\log 2}} + w_2 \left| \sqrt{NND(G)} - \sqrt{NND(G')} \right| + \frac{w_3}{2} \left(\sqrt{\frac{\mathcal{J}(P_{\alpha G}, P_{\alpha G'})}{\log 2}} + \frac{\mathcal{J}(P_{\alpha G^c}, P_{\alpha G^{c\prime}})}{\log 2} \right)$$

Where network node dispersion (NND) is a measure of the heterogeneity of graph *G* in terms of connectivity distances introduced by Schieber et al. (2017):

$$NND(G) = \frac{\mathcal{J}(\boldsymbol{P}_1, \dots, \boldsymbol{P}_N)}{\log(d+1)}$$

$P_i(j)$ being the fraction of nodes that are connected to node *i* at distance *j*, leading to the following equations:

$$\boldsymbol{P}_i = \{p_i(j)\};$$

$$\mathcal{J}(\boldsymbol{P}_1, \dots, \boldsymbol{P}_N) = \frac{1}{N} \sum_{i,j} p_i(j) \log(\frac{p_i(j)}{\mu_j})$$

$$\mu_j = (\sum_{i=1}^{N} p_i(j))/N$$

Where w1, w2 and w3 are arbitrary weights of the terms where w1+w2+w3=1.

Through their experiments, the weights of the three components should ideally be set to 0.45, 0.45, and 1, respectively. Additionally, applying the D-measure to real networks significantly increases the cost of computing the α-centrality of their graph's complements due to the sparse nature of real networks. Therefore, a simplified D-measure was used in this study, considering only the first two terms of the equation to assess network dissimilarity. When two graphs are isomorphic, the D-measure returns 0; otherwise, it quantifies the structural differences between the graphs.

4.3.3 Geometry and topology co-location analysis

Considering the possibility that pipeline and road networks may not be perfectly aligned, as well as potential data discrepancies leading to network misplacement, network overlap in the co-location analysis is defined as elements that are mutually parallel within a certain geometric distance. Similarly, due to considerations of data quality and coverage, road networks, rather than pipeline networks, were selected for buffer generation. A pipeline is considered to overlap with a road if it is located within the buffer of that road.

Buffer width reflects the stringency of the co-location analysis, where increasing the buffer size not only makes it more likely to capture pipelines overlapping with roads but also inevitably increases the inclusion of pipeline segments that do not overlap with roads, such as those crossing streets. Therefore, a sensitivity analysis was conducted to determine the appropriate buffer width by examining the pipeline length and proportion covered by buffers of different sizes (in 1-metre increments). Since the degree of alignment between pipelines and roads is unknown, the assumed buffer width was based on the common lane widths and lane numbers provided by the road classification frameworks from Kigali and Kisumu (see Tables 4.2 and 4.3). Alternative buffer distances (i.e., assumed street widths) ranged from 1 to 15 metres for Kisumu roads and 1 to 17 metres for Kigali roads, following methods outlined by Klinkhamer et al. (2017), Mair et al. (2017) and Klinkhamer et al. (2019). A maximum buffer width of 15 metres for Kisumu corresponds to the width of a dual-lane road, according to the Kenya Road Design Manual (Ministry of Roads and Transport, 2023). Meanwhile, 17 meters corresponds to the maximum road width of high-capacity urban roads in Kigali (City of Kigali, 2020c). By calculating the proportion of the pipeline network covered by buffers of varying widths, appropriate buffer widths were selected for the co-location analysis. The underlying logic is that if increasing the buffer width captures more pipeline fragments running parallel to roads than those crossing streets, the covered pipeline length should increase rapidly; otherwise, the rate of increase should slow down. The second derivative of the proportion of overlapping pipelines as the buffer width increases was used to identify these two growth trends.

Table 4.2 Kenya road classification and width (Ministry of Roads and Transport, 2023)

Road category	Most common functional class and type	Surface type	Lane width range (m)
Inter-urban roads	A (international highways)	Paved	Dual 2*7.3
	B (national highways)	Paved	7.3
	C (primary roads/inter county roads)	Paved	7.0
Rural roads	D (secondary roads/inter sub-county roads)	Paved	7.0
	E (minor roads/ sub-county roads)	Paved or unpaved	6.5 - 7.0
	F, G, P, S, W, T, U (local roads)	Paved or unpaved	5.5 - 7.0
Urban roads	UA (urban arterial roads)	Paved	Dual 2*7.3
	UC (urban collector roads)	Paved	7.3
	UL (local urban streets)	Paved	6.0 - 7.0

Table 4.3 Rwanda road classification and width (City of Kigali, 2020c).

Type	High-Capacity Urban Roads	Major Arterial Roads				Minor Arterial Roads			Collector Roads	
	Trunk Roads		Bus Rapid Transport	Link Roads	CBD Throughfare		Bus Routes	Commercial Streets	Residential Streets	Rural Road
Desirable Road Reserve Width	37 – 44 m	34 – 37 m	34-40m	34 – 37 m	28 – 37 m	22 – 27 m	27 m	27 m	18 – 22 m	18 – 22 m
Typical number of lanes per direction	2 – 5 lanes	2 – 4 lanes	2 – 3 lanes	2 – 3 lanes	2 – 3 lanes	1 – 2 lanes	1 – 2 lanes	1 – 2 lanes	1 – 2 lanes	1 – 2 lanes
Minimum Carriageway Width	3.5 m per lane	3.5 m per lane	3.5 m per lane	3.5 m per lane	3.5 m per lane	3.5 m per lane	3.5 m per lane	3.5 m per lane	3 m per lane	4 m per lane
Median Width	4 m	1 – 4 m	1 – 4 m	0.6 – 4 m	0.6 m	0.6 m	0.6 m	0.6 – 2 m	-	-
Hard Shoulder	3 m	-	-	-	-	-	-	-	-	-
Easement / Verge	2.5 – 6 m	2.5 – 6 m	-	-	-	-	-	-	-	2 – 3.5 m
Footway	-	-	1.5 m min	1.5 m min	1.5 m min	1.5 m min	2 m min	2 m min	1.5 m min	-
Cycleway	-	-	1.5 m min	1.5 m min	1.5 m min	1.5 m min	1.5 m min, or omit	1.5 m min, or omit	1.5 m min	-
Planting Strip	-	-	2 m	2 m	2 m	2 m	2 m	2 m	2 m	-

4.4 Results

The results are analysed from three perspectives to examine the relationship between urban morphology and pipeline networks. First, an overview of the topological properties of the networks in each city is provided, offering insights into their commonalities and, by extension, the influence of the urban environment on infrastructure network structures. Based on the same dual graphs, network dissimilarity was also calculated for networks within each city. Subsequently, different buffer widths were tested against the length of piped infrastructure elements that overlapped with roads, and the buffer widths that best captured network overlap patterns were selected for co-location analysis. The co-location analysis extracted overlapping urban networks and compared the centrality of overlapping elements in pairs, enabling a spatial and morphological comparison of network differences.

4.4.1 Topological characteristics of study infrastructure networks

Table 4.4 shows that the Kisumu sewer network exhibits a simple structure, characterised by high betweenness and closeness metrics, along with a relatively high average shortest path length. This simplicity is associated with lower overall connectivity and limited stability.

The table also reveals that the road networks in both Kisumu and Kigali exhibit lower link density. Moreover, the road networks demonstrate significantly lower betweenness and closeness centrality, along with higher average shortest path lengths compared to other networks. The clustering coefficient and meshedness are both indicators of faces or loops in a network, but they show opposing trends. The former identifies triangular loops, as discussed in literature review 2.4.1, which are typically rare in urban networks, while meshedness indicates that road networks tend to contain more square loops. This suggests that, in both cities, road networks are inherently less efficient than piped networks in terms of connectivity but exhibit greater robustness.

An analysis of network degree further supports this observation (Figures 4.8 and 4.9). In both networks, the water network exhibits a right-skewed distribution, with a few high-degree nodes (corresponding to edges in the network). This indicates the presence of hubs within the network, enhancing connectivity efficiency while reducing the water network’s resilience to risks.

Table 4.4 Summary of graph theory metrics for the dual representation of road, water, and sewer networks in Kisumu and Kigali

	Kisumu			Kigali	
	Water	Road	Sewer	Water	Road
Number of nodes	4076	11616	348	6721	35191
Number of edges	6690	22203	547	10781	69839
Link Density	0.0008	0.0003	0.0091	0.0005	0.0001
Mean Degree Centrality	3.2826	3.8228	3.1437	3.2082	3.96914
Mean Betweenness Centrality	0.0092	0.0044	0.0447	0.0054	0.0020
Mean Closeness Centrality	0.0275	0.0200	0.0630	0.0276	0.0141
Cluster Coefficient	0.4944	0.4006	0.4571	0.4307	0.3909
Meshedness Coefficient	0.3210	0.4558	0.2894	0.3022	0.4923
Average Shortest Path Length	38.541	52.576	16.478	37.546	71.769

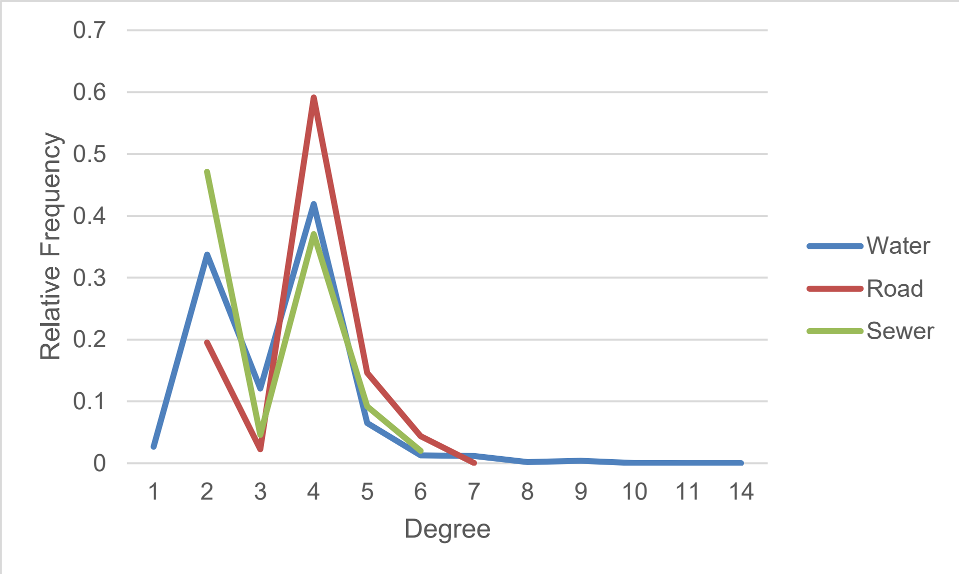


Figure 4.8 Degree distribution of the Kisumu water pipeline, road, and sewer pipeline networks.

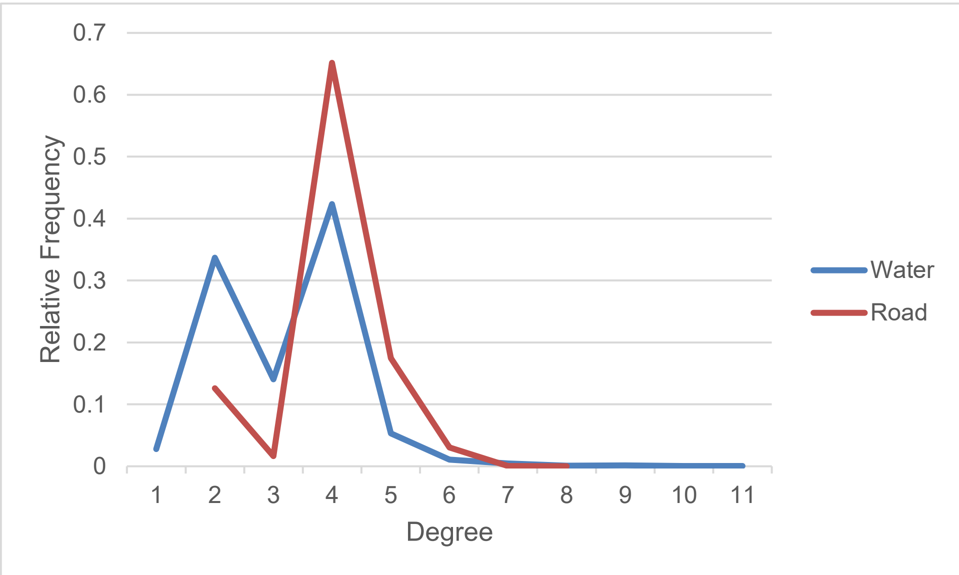


Figure 4.9 Degree distribution of the Kigali water pipeline and road networks.

Using the D-measure, pairwise dissimilarity comparisons of infrastructure networks from both study cities were conducted. Since a value of 0 indicates two isomorphic networks, while larger values signify greater structural differences, the results reveal several key patterns (Table 4.5). First, infrastructure networks generally all exhibit high structural similarity. Second, the similarity between Kisumu's road and water networks is higher than that observed between Kigali's road and water networks. Third, Kisumu's sewer network shows markedly high dissimilarity compared to other networks. Finally, networks of the same type display the highest similarity: both the water and road networks in Kigali and Kisumu exhibit relatively low dissimilarity, with the similarity between the two cities' water networks being particularly pronounced. To further check this finding, the study additionally compared the dissimilarity between Kisumu's pipeline networks and Kigali's road network, as well as between Kigali's pipeline networks and Kisumu's road network. The results suggest that, broadly speaking, infrastructure networks within the same city are more structurally similar to one another than to networks of the same type in another city.

Table 4.5 Dissimilarity index between road, water, and sewer networks

	Kisumu Water	Kisumu Road	Kisumu Sewer	Kigali Water	Kigali Road
Kisumu Water	0				
Kisumu Road	0.146	0			
Kisumu Sewer	0.290	0.357	0		
Kigali Water	0.072	0.161	0.315	0	
Kigali Road	0.277	0.152	0.427	0.284	0

4.4.2 Sensitivity analysis of buffer size effects on network co-location

The results of the sensitivity test (Figure 4.10) indicate that as the buffer size increases, the increase in the length of water and sewer pipelines covered by roads follows a pattern similar to a logarithmic distribution. The second derivative results (Figure 4.11) further show that the trends in both cities can be divided into three phases, corresponding to rapid, moderate, and slow increases in the proportion of covered pipeline length. The point at which the increase slows down occurs at a buffer width of 10m in Kisumu, while in Kigali, it occurs at 11 meters. Accordingly, the buffer width for co-location analysis of Kisumu roads was set at 10 meters, aligning with the standard road width (including shoulders) of approximately 10 meters in Kenya (Ministry of Roads and Transport, 2023). In this case, the covered length of the water pipeline is 371,496.4 meters, constituting 69.624% of the entire network length. The covered length of the sewer pipeline is 52,888.12 meters, representing 58.0246% of the total sewer network length. Meanwhile, in Kigali, with a buffer width of 11 meters, the road network overlaps with 1,378,317.75 meters of water pipelines, accounting for 79.47% of the total pipeline network length.

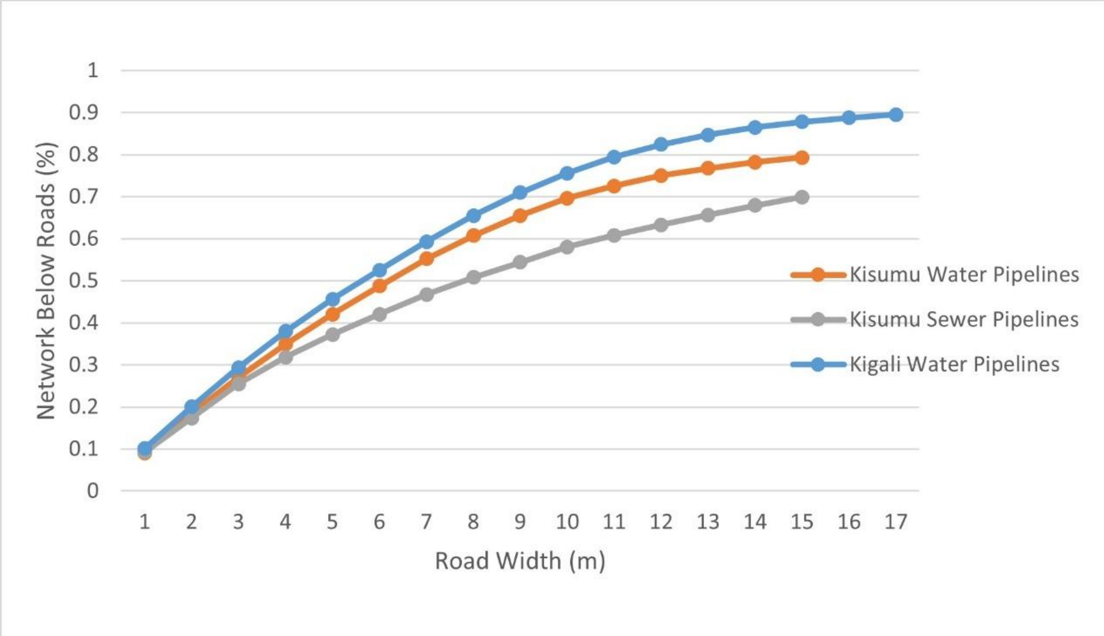
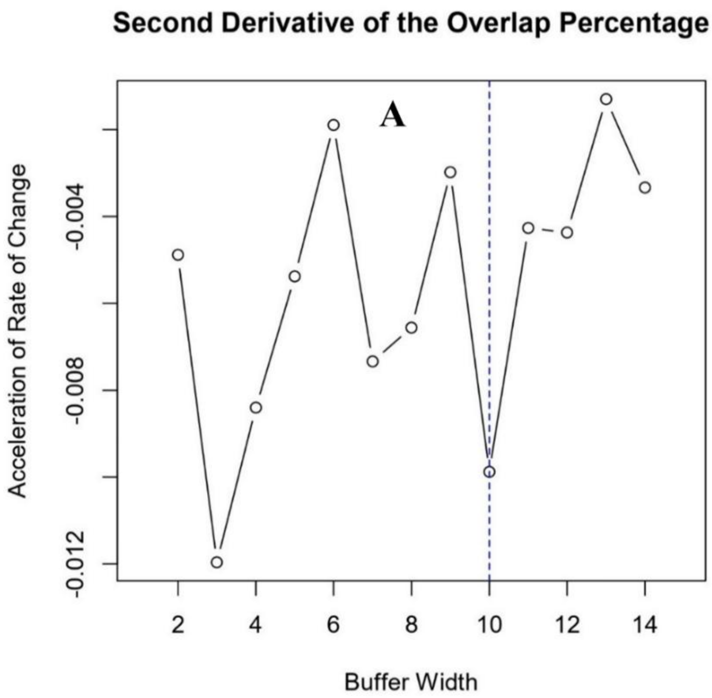


Figure 4.10 Sensitivity analysis of additional street width and covered network length below the roads.



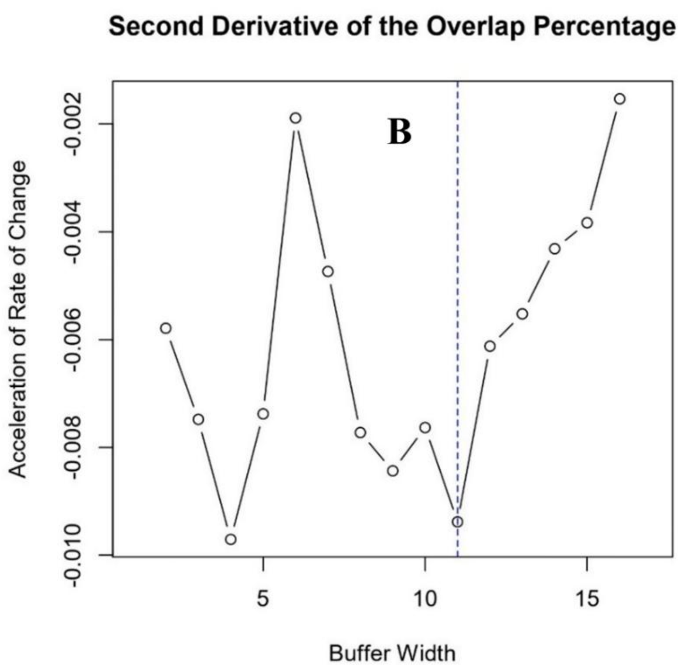


Figure 4.11 Second derivative distribution of the proportion of pipeline length covered by road buffers at different widths. Panel A shows Kisumu, where the proportions of water and sewer pipelines were averaged in the analysis. Panel B shows Kigali, focusing solely on the water pipelines.

4.4.3 Co-location of infrastructure networks

The initial geographic visualization of network degree centrality (Figures 4.12 and 4.13) reveals that high-degree edges appear in all three networks, extending beyond the central region. In both cities, road networks generally exhibit a decreasing degree trend from the core outward. Notably, Kigali's road network features multiple high-degree clusters, as reflected in the degree distribution (Figure 4.9), whereas the pipeline network does not exhibit a similar pattern. This observation aligns with the structural differences highlighted by the graph metrics, which indicate that pipeline networks have a more evenly distributed structure compared to road networks.

The maps also provide insight into potential spatial correlation of high-degree distribution. For instance, in Kisumu, water pipelines exhibit relatively high degree within slums (see Section 3.2.2), a phenomenon not observed in the road network. Conversely, in regions where road degree is high, pipelines do not display a similarly high-degree distribution. However, high-degree sewer pipelines tend to co-locate with high-degree roads.

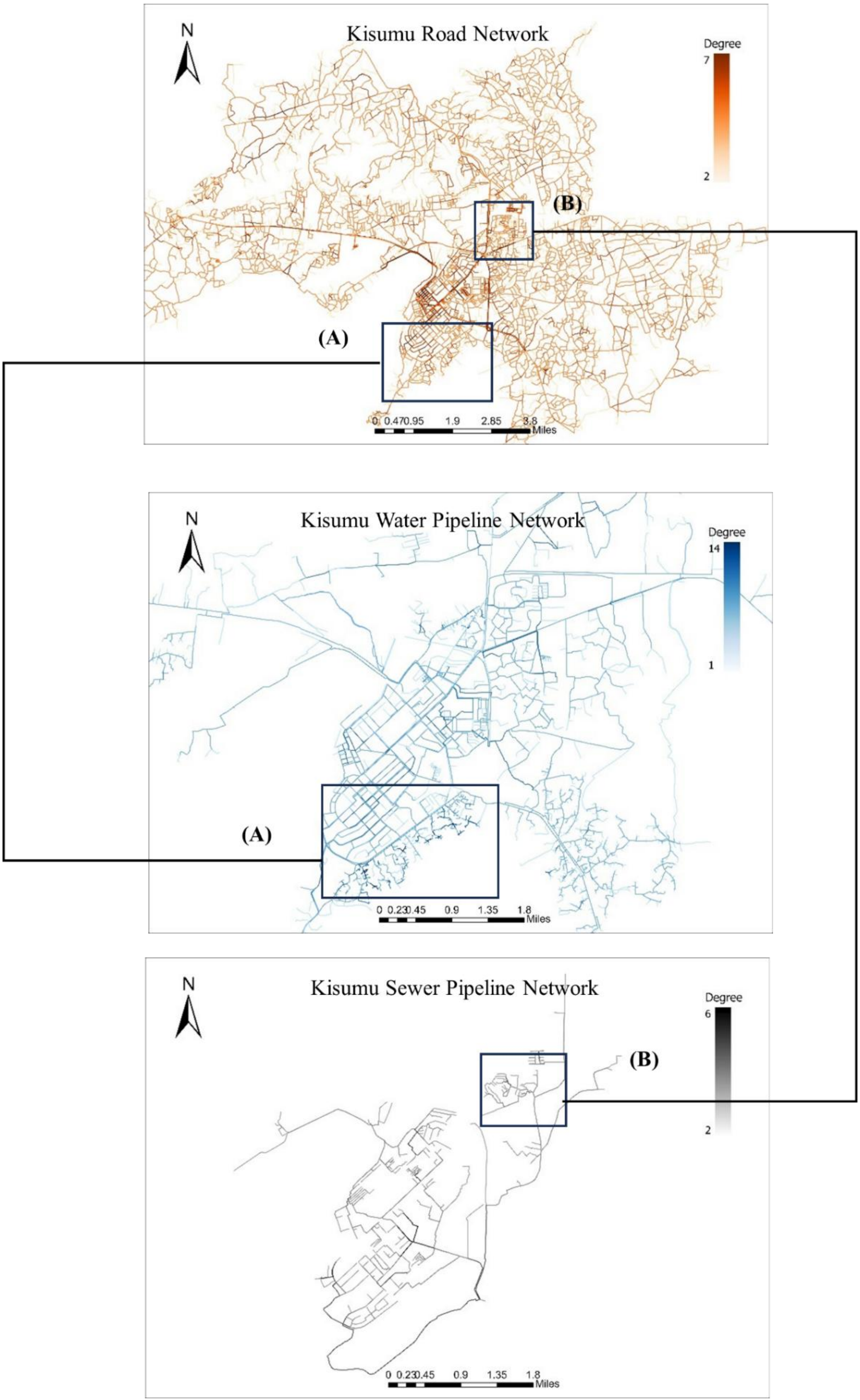
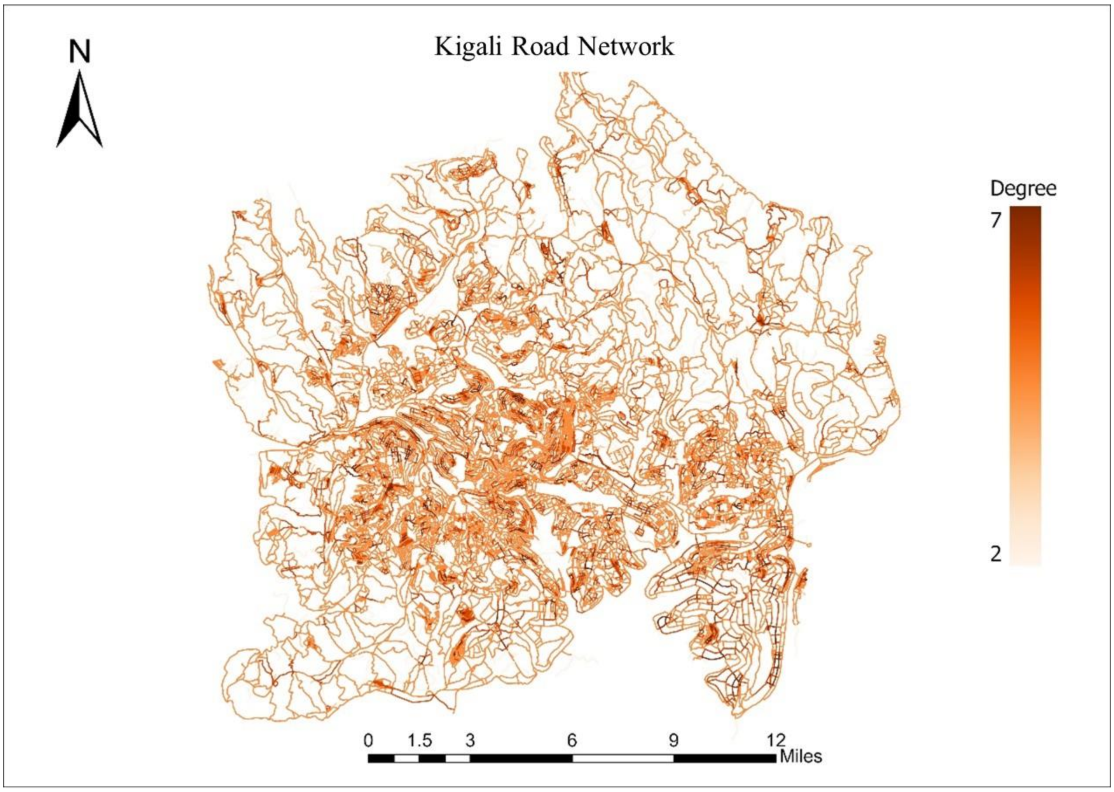


Figure 4.12 Remapping of the dual graph degree to the primal space for the Kisumu road network, water pipeline network, and sewer pipeline network of Kisumu. The box highlights the different distributions of high-degree roads and pipelines within the same area (A), while showing similar distributions between roads and the sewer pipeline network (B).



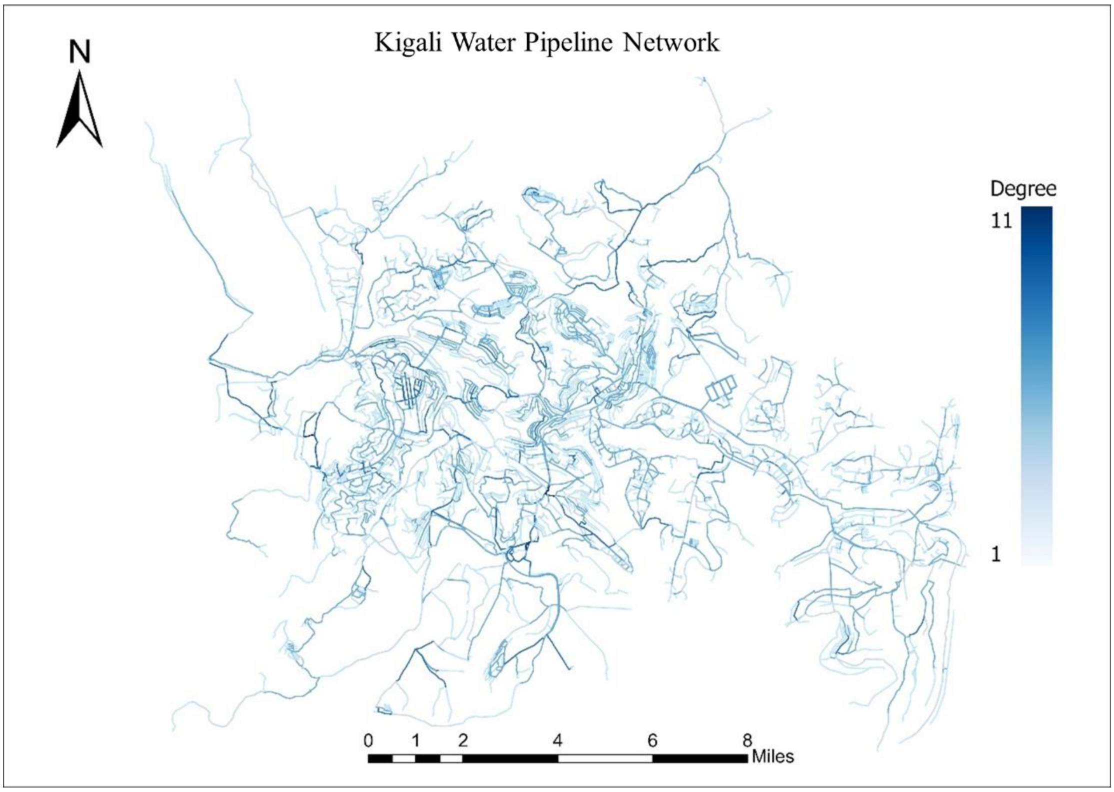


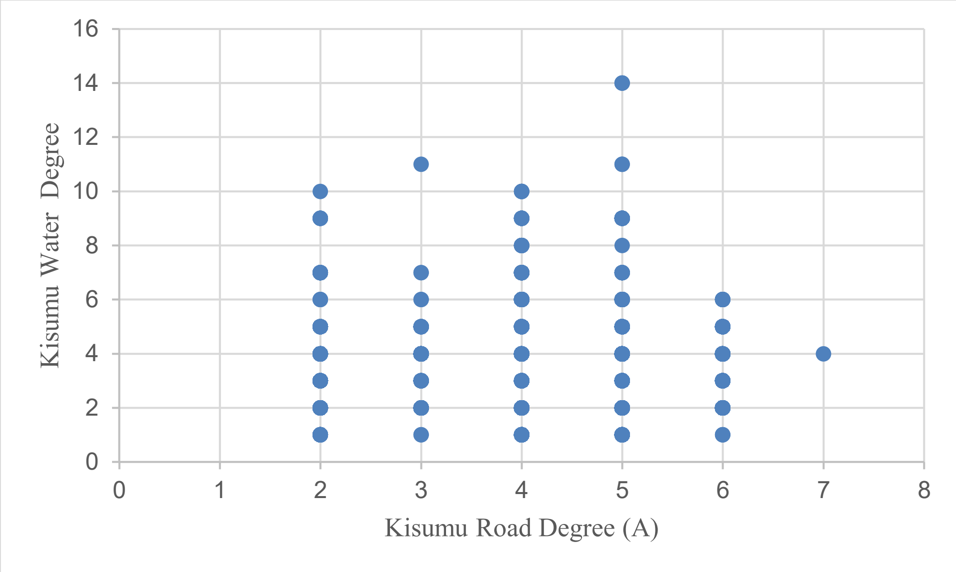
Figure 4.13 Remapping of the dual graph degree to the primal space for the Kigali road network and water pipeline network. The two panels show the overall distribution differences of high-degree roads and pipelines across the area.

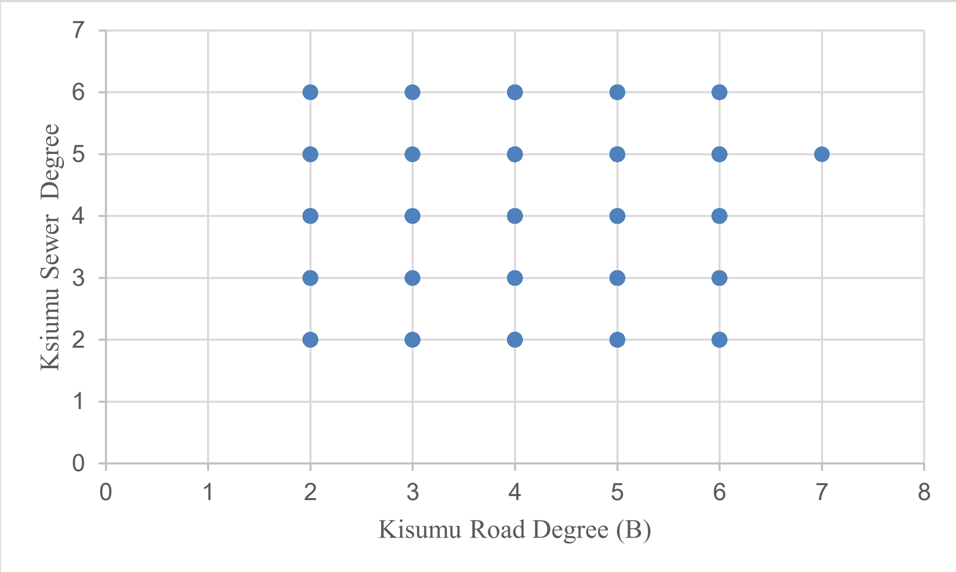
We further examined centrality by applying a sensitivity test to extract pipelines overlapping with roads using the buffers (Figures 4.14 and 4.15). By analysing the degree, betweenness, and closeness centrality distributions for co-located features across the five infrastructure networks, several patterns emerge:

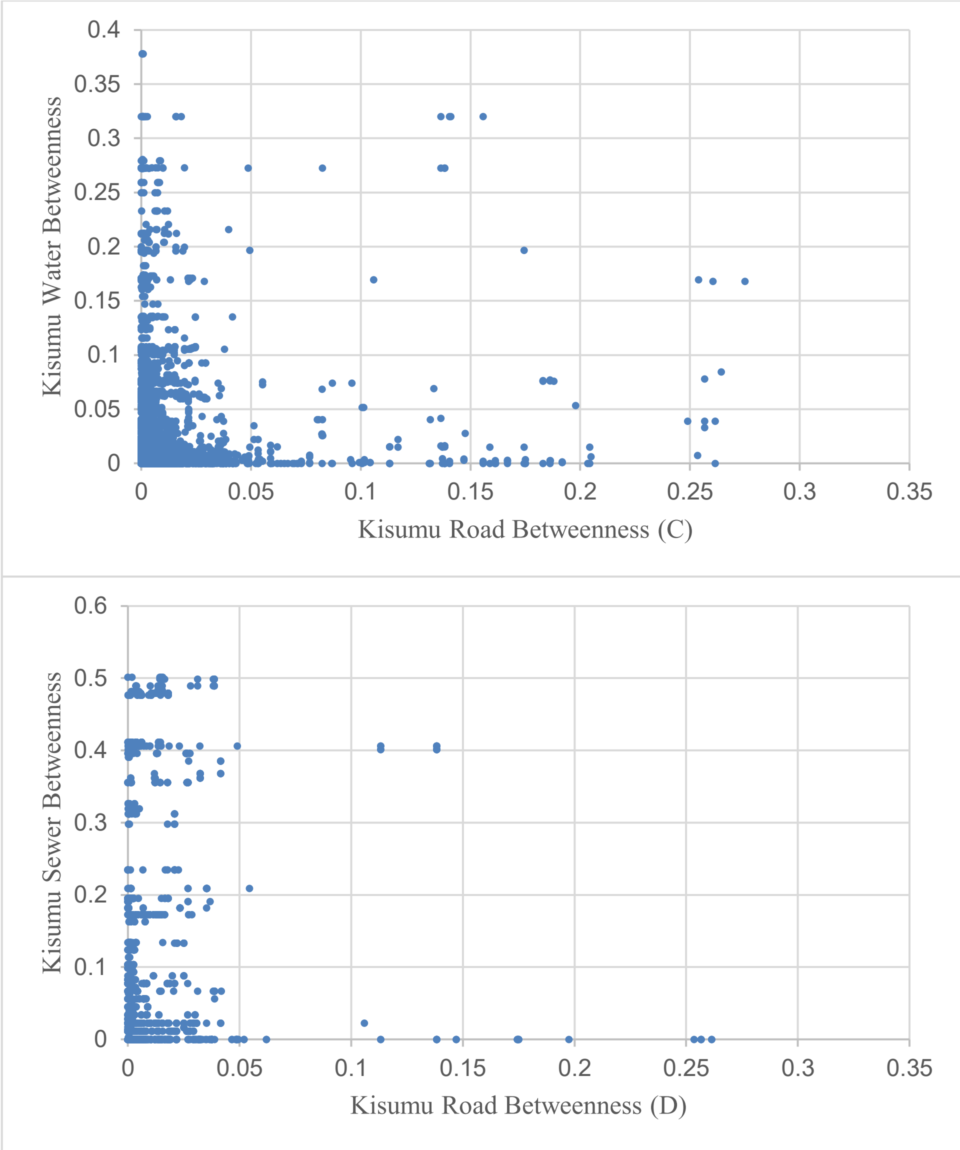
Degree distribution: In both cities, the degree distributions of the networks are generally uniform (Figures 4.14 (A)(B), Figure 4.15 (A)). There is no significant dependency between the degrees of pipelines and roads. However, high-degree roads tend to overlap with low-degree pipelines (as shown in the example area in Figure 4.12 (A)), a trend that is slightly less pronounced in sewer pipelines (Figure 4.12 (B)).

Betweenness centrality: High-betweenness edges in one network tend to overlap with low-betweenness edges in the other (Figure 4.14 (C)(D), Figure 4.15 (B)). However, in Kigali, some high-betweenness pipelines overlap with roads exhibiting a wider range of betweenness values.

Closeness centrality: There is no correlation between the closeness centrality of roads and sewers, suggesting a lack of significant planning coordination between the two networks (Figure 4.14 (F)). However, in both cities, water pipelines exhibit an approximately linear relationship with road closeness (Figures 4.14 (E)(F), Figure 4.15 (C)), meaning that high-closeness roads tend to overlap with high-closeness pipelines.


Kisumu Water Degree
Kisumu Road Degree (A)
0
1
2
3
4
5
6
7
8
0
2
4
6
8
10
12
14
16


Ksiumu Sewer Degree
Kisumu Road Degree (B)
0
1
2
3
4
5
6
7
8
0
1
2
3
4
5
6
7


Kisumu Water Betweenness
Kisumu Road Betweenness (C)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
Kisumu Sewer Betweenness
Kisumu Road Betweenness (D)
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35

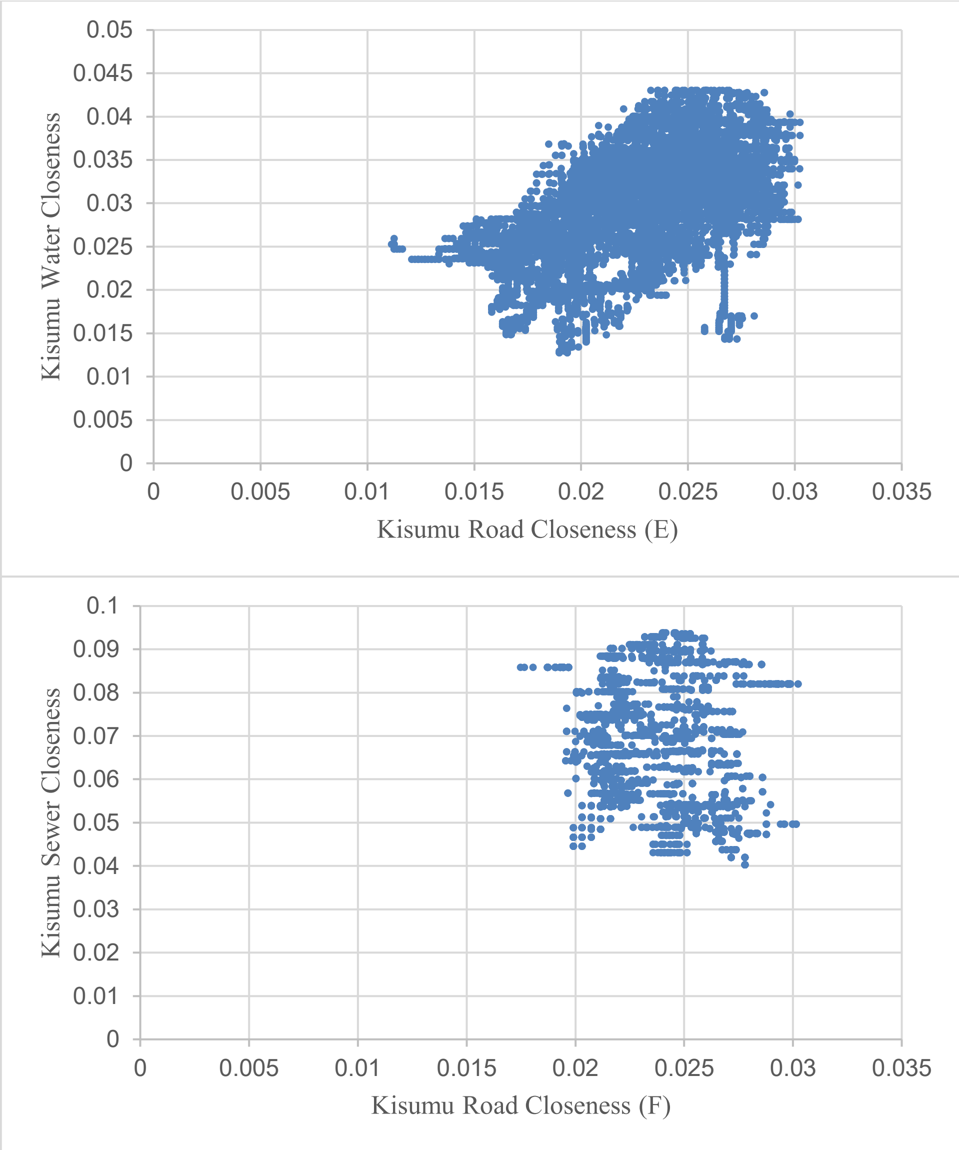


Figure 4.14 Dual mapping local graph metrics (node degree(A and B), betweenness(C and D), and closeness(E and F)) of co-located roads, water and sewer pipelines in Kisumu.

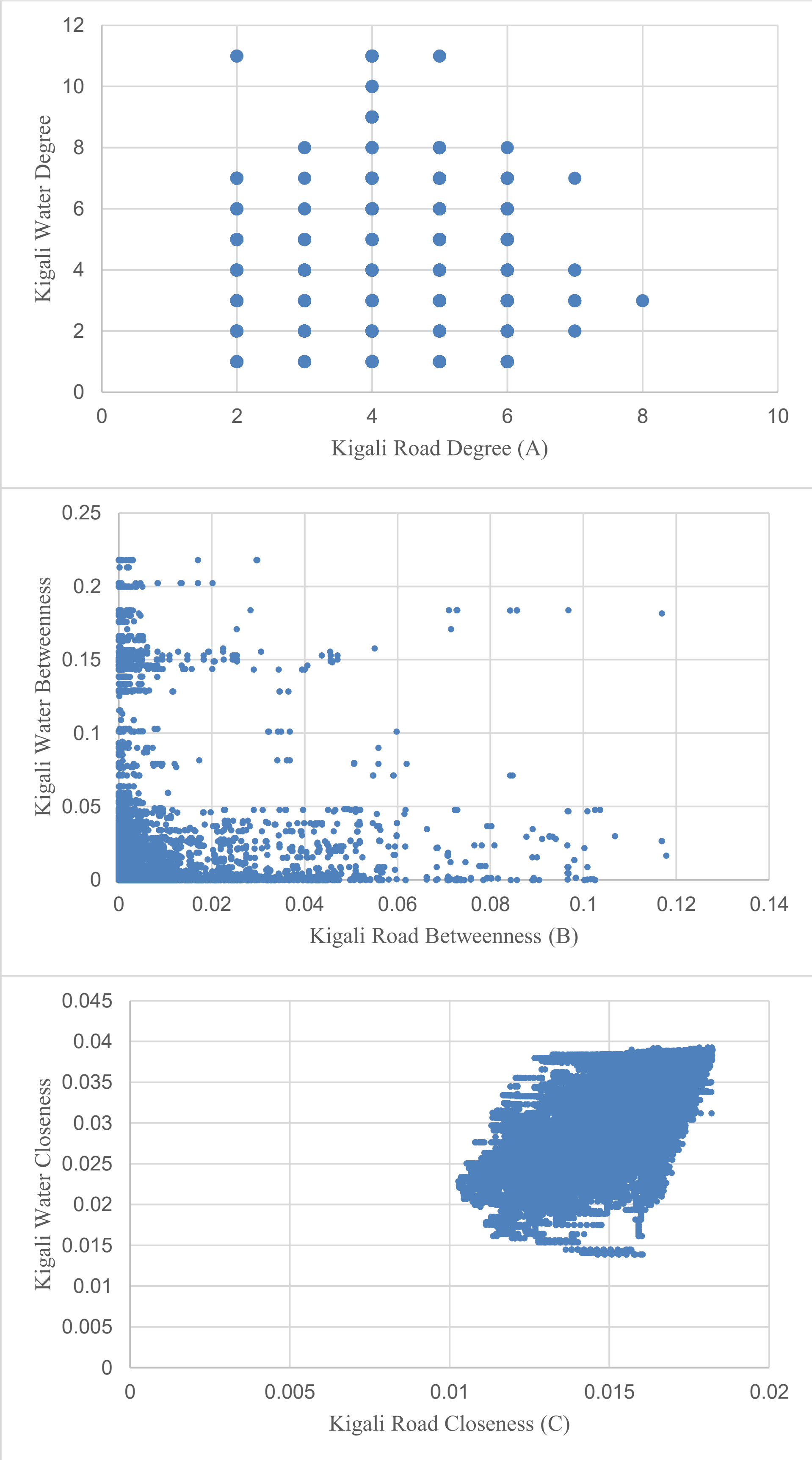


Figure 4.15 Dual mapping local graph metrics (node degree (A), betweenness (B), and closeness (C)) of co-located roads and water pipelines in Kigali.

4.5 Discussion

Urban infrastructure networks are shaped by urban morphology, leading to commonalities in network structure across cities of similar scale. The study posits that SSA cities, facing comparable urbanisation challenges, exhibit shared characteristics in their infrastructure networks. These commonalities can be identified by examining the structural similarities between road networks and water and sanitation networks. To assess them from both topological and geometric perspectives, the study applied a dual transformation to the networks, measured their graph properties, calculated network distances, and analysed the centrality metrics of pipelines co-located with roads.

We find that despite differences in scale and national context, infrastructure networks in SSA cities exhibit remarkable commonalities, even though the networks in the two cities were developed independently. Notably, infrastructure networks, including water pipelines, share a high degree of similarity with road networks. Graph metrics reveal a consistent pattern in which road networks tend to be more structurally robust compared to water networks. More specifically, networks of the same infrastructure type exhibit higher similarity, as demonstrated by the D-measure. Moreover, infrastructure networks within the same city tend to be more similar to each other, while differences between roads and water networks in Kisumu and Kigali also follow similar patterns. The centrality analysis of co-located networks further reinforces this observation, as the distribution trends of centrality measures in co-located elements are strikingly similar across both cities.

In studies of African cities, urbanization patterns are generally characterized by a concentration of population in the central business district (CBD), with a rapid decline in transport network density and accessibility towards the urban periphery, in contrast to European cities (Antos et al., 2016). This observation aligns with the study's graph analysis (Table 4.4) and degree distribution maps (Figures 4.12 and 4.13), which show that in both Kisumu and Kigali, the degree of dual road networks decreases outward from the centre, accompanied by relatively low link density. A critical factor shaping SSA urban morphology and network structure is informality, which introduces heterogeneity in both spatial patterns and network properties. Studies on urban form have identified two primary typologies: organically developed road networks, shaped by historical and geographical factors, exhibiting irregular patterns, and planned road networks, which follow structured, grid-like patterns due to formal urban planning interventions. For instance, Wang (2015) compared the road networks of Beijing and London, revealing the contrast between Beijing's top-down planned grid structure and London's self-organized tree-like network. Similarly, Alobaydi and Rashid (2024) classified Baghdad's urban form into six categories based on organic and grid-based structures, each shaped by different historical periods. Comparable studies on African cities have also categorized urban forms into irregular and well- structured patterns. Baffoe et al. (2020b) classified Kigali's neighbourhoods into formal, informal, and mixed types based on their planning characteristics, while Steyn (2012) highlighted the morphological disparities between slums and formal settlements in Kisumu. The predominant urban form inevitably influences overall street network properties (Serra et al., 2016). This relationship was examined through graph analysis. The network metrics (Table 4.4) and degree distribution plots (Figures 4.8 and 4.9) indicate that the degree of road intersections in both cities is close to four. In the dual graph representation, degree corresponds to the number of intersections along a road. A grid-based layout would typically yield higher degree values due to the increased number of dual connections. For example, Porta et al. (2006) reported degree values of 8 and 6 for street samples in San Francisco and Barcelona, respectively, while Wang (2015) found that the average degree in the dual graph of Beijing (a grid-based city) was 5, compared to 3.6 in London (a self-organized city). Based on this comparison, the road networks of Kisumu and Kigali exhibit characteristics more aligned with self-organized structures. The spatial distribution of high-degree roads, as visualized in the degree maps (Figures 4.12 and 4.13), further supports this distinction. In Kisumu, high-degree roads are concentrated in formal areas and a few well-planned slums, while in Kigali, the pipeline network displays widespread clustering of high-degree nodes. These cluster locations align with household living condition analyses by Akinyemi and Bigirimana (2012) and the community survey conducted by Baffoe et al. (2020b). These graph indicators also align with existing knowledge on the distribution of slums in both cities. In Kisumu, slums vary in their degree of planning—Manyatta, for instance, has a more structured road network compared to Obunga and Bandani (UN-Habitat, 2005). In contrast, reports on Kigali indicate that regular residential land is consistently distributed throughout the city (Antos et al., 2016, City of Kigali, 2020c).

The similarity between road morphology and pipeline networks highlights the influence of urban morphology on infrastructure development. Previous studies have shown correlations between road networks and pipeline networks, with degree distributions of roads, urban drainage networks, and water distribution networks fitting Pareto distributions (Klinkhamer et al., 2019). Based on an analysis of urban drainage networks, the authors argue that urban networks evolve through preferential attachment while being constrained by factors such as cost and available space. Building on these insights, this study identifies differences in the graph-theoretic properties of road and infrastructure networks within the dual space in Kisumu and Kigali. These differences manifest both in global performance and spatial distribution of graph attributes. In Kisumu, the graph indicators suggest that roads exhibit greater structural stability than water networks, which in turn are more stable than sewer networks. Conversely, global centrality measures indicate that network efficiency follows the opposite trend. A similar pattern is observed in Kigali, suggesting that these differences may stem from variations in urban network development principles. Studies on road network evolution indicate that urbanisation involves two predominant processes—densification and exploration—which dominate at different stages of city development (Gudmundsson and Mohajeri, 2013, Mohajeri and Gudmundsson, 2014, Strano et al., 2012). Over time, urban networks tend to mature and transition towards more structured grid patterns (Strano et al., 2012). However, pipeline networks face additional constraints, leading to greater variability in their construction. Ozanne (2011) notes that the primary determinant in pipeline network construction is the shortest path between origin and destination, as cost considerations play a crucial role. However, additional factors such as land ownership and environmental constraints (e.g. terrain and soil conditions) must also be accounted for. These factors are inherently linked to urban morphology. First, population distribution influences pipeline terminal locations, and population patterns are recognised as key drivers of urban expansion (Achibet et al., 2014), which in turn affects the distribution and density of buildings (Prieto-Curiel et al., 2023). Second, slums often exhibit insecure land tenure and tend to develop in steep or marginal areas (McCartney and Krishnamurthy, 2018). This is particularly evident in Kigali, where, due to its topography, slums are predominantly located on steep hillsides and marshy lowlands (Manirakiza et al., 2019).

Spatially, there are differences in the centrality distribution of infrastructure networks between Kisumu and Kigali. The degree maps (Figures 4.12 and 4.13) indicate two possible scenarios:

Densely developed areas with intensive infrastructure:
The co-location of high-degree edges is likely prevalent in areas characterised by a concentration of infrastructure. These regions typically exhibit significant urban development and a high demand for multiple types of infrastructure.

Critical arterials of urban connectivity:
Co-located high-degree edges may align with key urban arteries, which serve as major transportation or utility routes. These critical corridors play a fundamental role in the overall functionality and connectivity of the city.

Similarly, the colocation analysis of betweenness centrality for roads and pipelines reveals notable differences. One contributing factor is the disparity in the spatial coverage of road and pipeline networks, meaning that well-developed pipeline systems do not necessarily correspond to the communities connected by road networks. This suggests a potential imbalance in the development of road and pipeline infrastructure in both cities, which correlates with neighbourhood distribution. In Kigali, certain high-betweenness pipelines overlap with roads exhibiting diverse betweenness values. Conversely, the absence of correlation between road network closeness centrality and sewer networks indicates a lack of significant planning coordination between these two systems. However, water pipelines exhibit an approximately linear relationship with road network closeness in both cities, implying that roads with high closeness centrality tend to overlap with water pipelines of similarly high closeness. This finding underscores the influence of urban accessibility on the spatial organisation of essential infrastructure.

In contemporary urban network research, substantial attention has been given to infrastructure studies. However, there remains a gap in understanding the relationship between urban morphology and infrastructure beyond road networks, particularly in the context of SSA urbanisation. This study contributes to filling gap by employing graph analysis of urban road networks to characterise SSA urban morphology. The study first proposes a workflow to address data quality issues posed by data deficiencies in SSA countries, including strategies for topology correction, graph smoothing, and simplification. Using methods such as D-measure and co-location analysis, the study provides empirical evidence on how the unique urban forms of SSA cities influence infrastructure development. Additionally, this analysis offers insights into the issue of urban scaling, which concerns the variation of urban indicators across cities of different sizes (Pumain and Guerois, 2004). The findings reveal a relatively stable relationship between road and pipeline networks in Kisumu and Kigali, despite their differing scales. This observation aligns with previous research on human interaction networks(Schläpfer et al., 2014). It is important to note that, beyond the D-measure used in this study, various approaches exist for assessing graph similarity (i.e., graph isomorphism). For instance, information-theoretic methods define graph similarity through information compression (Coupette and Vreeken, 2021). Meanwhile, Graph Neural Network (GNN) (Gori et al., 2005) and Graph Convolutional Network (GCN) approaches also assess graph similarity in the graph signal processing way, a framework that treats network features as signals (Dong et al., 2020, Ma et al., 2021, Ortega et al., 2018). Furthermore, Kolowa et al. (2024) highlight the association between unplanned, low-density sprawl—a characteristic of SSA cities—and street accessibility. By evaluating network connectivity, they find that the presence of informal settlements does not always correlate with urban sprawl in SSA. This suggests that assessing the impact of slums, a defining feature of SSA urbanisation, on pipeline networks requires further investigation. This topic is explored in greater depth in the next chapter.

4.6 Conclusion

Urban morphology has a significant role in shaping infrastructure networks in SSA cities. Through graph analysis, the study demonstrate the structural relationships between road and pipeline networks in Kisumu and Kigali, revealing the characteristics of infrastructure in SSA cities and their relationship with urban morphology. The application of graph metrics provides empirical evidence on how SSA's unique urban forms influence infrastructure distribution. Furthermore, the results of D-measure and colocation analysis indicate that while road and pipeline networks exhibit a relatively stable relationship across cities of different scales, variations in their connectivity and coverage suggest imbalances in infrastructure development.

Chapter 5 An analysis of pipeline network topology and urban environment in Kisumu and Kigali

This chapter forms the basis of a paper accepted for publication as:
'Graph and Community Detection Analysis of Pipeline Network Configuration and Urban Morphology in Kisumu and Kigali,' accepted for publication in the *Journal of Water Resources Planning and Management*.

The Sustainable Development Goal (SDG) 6 incorporates safely managed drinking water as an indicator, defined as on-premise improved water sources, including piped water (WHO/UNICEF, 2018). Achieving this goal necessitates specific requirements for the infrastructure of urban piped systems. As urban populations grow, service demand often outpaces infrastructure development (Adams et al., 2019). This disparity makes it challenging for formal water services to reach all residents, deviating from the ideal of networked cities. In informal settlements, two distinct piped water configurations exist. On one hand, the adverse locational factors of informal settlements, coupled with a lack of spatial planning, lead to chaotic pipe layouts, which increase the difficulties of water provision, management, and maintenance (as discussed in Section 2.3). On the other hand, some governments have implemented interventions to enhance water services and pipeline infrastructure within slums. To assess the pipeline characteristics of slums in the contexts of Kisumu and Kigali, this research analyses two primary factors: 1. The impact of slum conditions on piped water supply, and 2. The effects of measures taken by Kisumu on pipeline configurations.

5.1 Motivation and objectives

Urbanisation in Sub-Saharan Africa (SSA) is largely attributed to the expansion of informal settlements, which often lack spatial planning and are located in disadvantaged areas. Previous research on SDG 11 (concerning the upgrading of slums) shows that the unplanned dense buildings within slums create a particular pattern of road networks, affecting residents' spatial access to services (Brelsford et al., 2018). For similar reasons to other services, the limited space within slums also constrains pipeline laying for water and sanitation. Pipelines in slums, therefore, are expected to have distinctive morphology, which impacts the delivery of water services, as discussed in the literature review (Section 2.3).

Given the obstacles faced by governments in improving urban water supply, a range of alternative service delivery models has been proposed for SSA, with community-based paradigms gaining attention in recent years (Adams et al., 2019, Dos Santos et al., 2017). In the case of delegated management models (DMM), for instance, this policy has been implemented in the slums of some cities, assigning the construction and operation of household connections to small-scale water providers (individuals or local entrepreneurs closely linked to the community). This approach is perceived as a potential solution for delivering sustainable water and sanitation services in slums (see Section 2.2). One of the anticipated improvements brought by these small-scale water providers is the enhancement of the spatial configuration of pipelines (World Bank, 2009). The distribution of pipelines significantly impacts the delivery of water services; however, the spatial configuration of pipelines in slums has not been studied previously.

As demonstrated in the previous chapter, graph methods offer advantages in analysing urban networks. However, a few studies have applied graph theory to the field of water, with most of these located in the USA (Hwang and Lansey, 2017, Porse and Lund, 2016, Yazdani and Jeffrey, 2010), the UK (Yazdani and Jeffrey, 2012a, Yazdani and Jeffrey, 2012b), Italy (Pagano et al., 2019) and other developed countries. Research on water pipeline networks in LMICs remains sparse. Existing research has not explored the use of graph-theoretic methods to analyse the impact of urban communities or policies on pipeline networks. This chapter therefore aims to:

1. Develop a workflow for evaluating the connectivity and resilience of pipeline networks via graph theory metrics within two data-sparse case study cities in SSA, namely Kisumu and Kigali.
2. Assess the utility of the InfoMap algorithm for detecting pipeline communities and informing urban water service planning in the study areas.
3. Interpret the differences in pipeline distribution within urban areas through graph analysis and community detection outputs, taking into account the DMM and distribution of slums.

5.2 Overview of methodology

The analysis process was divided into three parts (Figure 5.1), utilizing the primal graphs of water pipeline networks in Kisumu and Kigali constructed in the previous chapter. Since there was no specific analytical objective concerning edge relationships, the study retained the primal graph format in this chapter without performing further dual mapping, as this facilitates the interpretation of metrics. Additionally, the global graph metrics were applied to the primal graphs, and the distribution of their centrality metrics was summarised using commonly employed statistical distributions. Finally, the Infomap algorithm (Rosvall et al., 2009) was used to identify discrete communities within the Kisumu and Kigali pipeline networks based on their topology. The distribution of detected communities was then interpreted based on the pattern of slums. The communities in Kisumu were also examined based on the pipelines managed under DMM. In the first two steps, the networks of Kigali and Kisumu were treated as unweighted and undirected. Considering that the continuity of water supply is one of the standards defining safely managed water ("available when needed") as stated by WHO/UNICEF (2018), vulnerability weights were assigned to both sets of pipelines during the community detection step. The following sections will sequentially address the delineation of slums and pipelines managed under DMM, as well as the data preparation, analysis, and community detection steps.

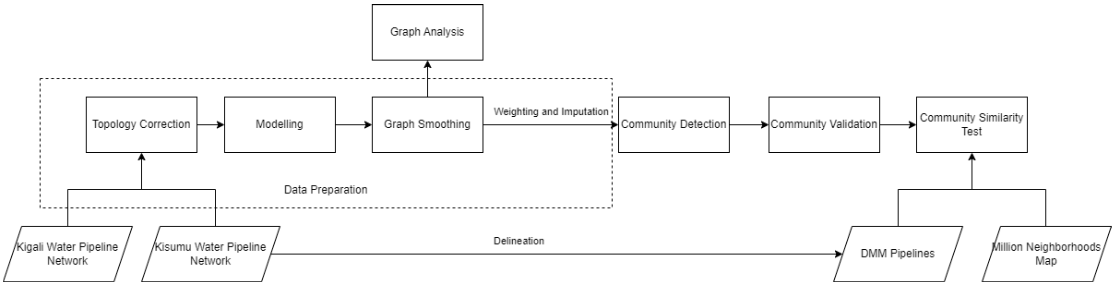


Figure 5.1 Flowchart of the major stages in graph analysis and community detection. Pipeline networks from Kisumu and Kigali were pre-processed for graph theory analysis and weighted for community detection. The community similarity test involved comparing the pipeline communities of Kigali and Kisumu with the Million Neighborhoods map, as well as a separate comparison of Kisumu's communities with the DMM information.

5.3 Data preparation

The pipeline networks of Kisumu and Kigali were selected to address the research objectives. To analyse the impact of slums and water management measures on these networks, it was first necessary to delineate both the pipelines within slums and those managed under the delegated management model. Additionally, considering concerns about network vulnerability, the material and diameter attributes of the pipelines were weighted and imputed for community detection.

5.3.1 Delineation of slums

As described in Chapter 4, Kisumu's water pipeline network is sourced from KIWASCO, while Kigali's network originates from the GIS for Water project. Both network data were topologically corrected and constructed to be converted into a simple primal graph prior to subsequent analysis.

To compare the morphological differences between pipelines in slums and other areas, it was first necessary to extract the pipelines within slum regions. Identifying slum areas is often challenging, with different studies adopting various approaches (Kuffer et al., 2016). In this study, the Million Neighborhoods map (Brelsford et al., 2018) provided under licence by the University of Chicago's Mansueto Institute for Urban Innovation was used (see Figures 5.2 and 5.3). Brelsford et al. (2018)’s Million Neighborhoods map assumes that topology, or spatial connectivity rather than geometry, determines the city's form. Since slums are characterised by unplanned spatial layouts, it is difficult for people living in slums to access services and infrastructure. Therefore, slums can be recognized through their access networks (roads, streets, and paths). The Million Neighborhoods map uses a metric called block complexity to measure the connectedness of a city block. In a dual-mapped graph where the internal parcels are represented by nodes, block complexity refers to the number of iterations required to continuously derive the dual of the graph until it converges into a trivial tree graph, reflecting the difficulty of travelling inside the block. As a result, the Million Neighborhoods map can be considered a slum map representing the complexity of the neighbourhoods (Chen et al., 2022b). Areas with higher complexity mean that street access from an area’s buildings is more difficult, a characteristic of slums. In Figure 5.2, areas with high block complexity in Kisumu generally decrease from south to north, while Figure 5.3 shows that in Kigali, they decrease from west to east. Additionally, a slum distribution map for Kisumu from the United Nations (UN-Habitat, 2005) and a land use map from the Kigali master plan (City of Kigali, 2020a) were used to help identify slums in Kisumu and Kigali.

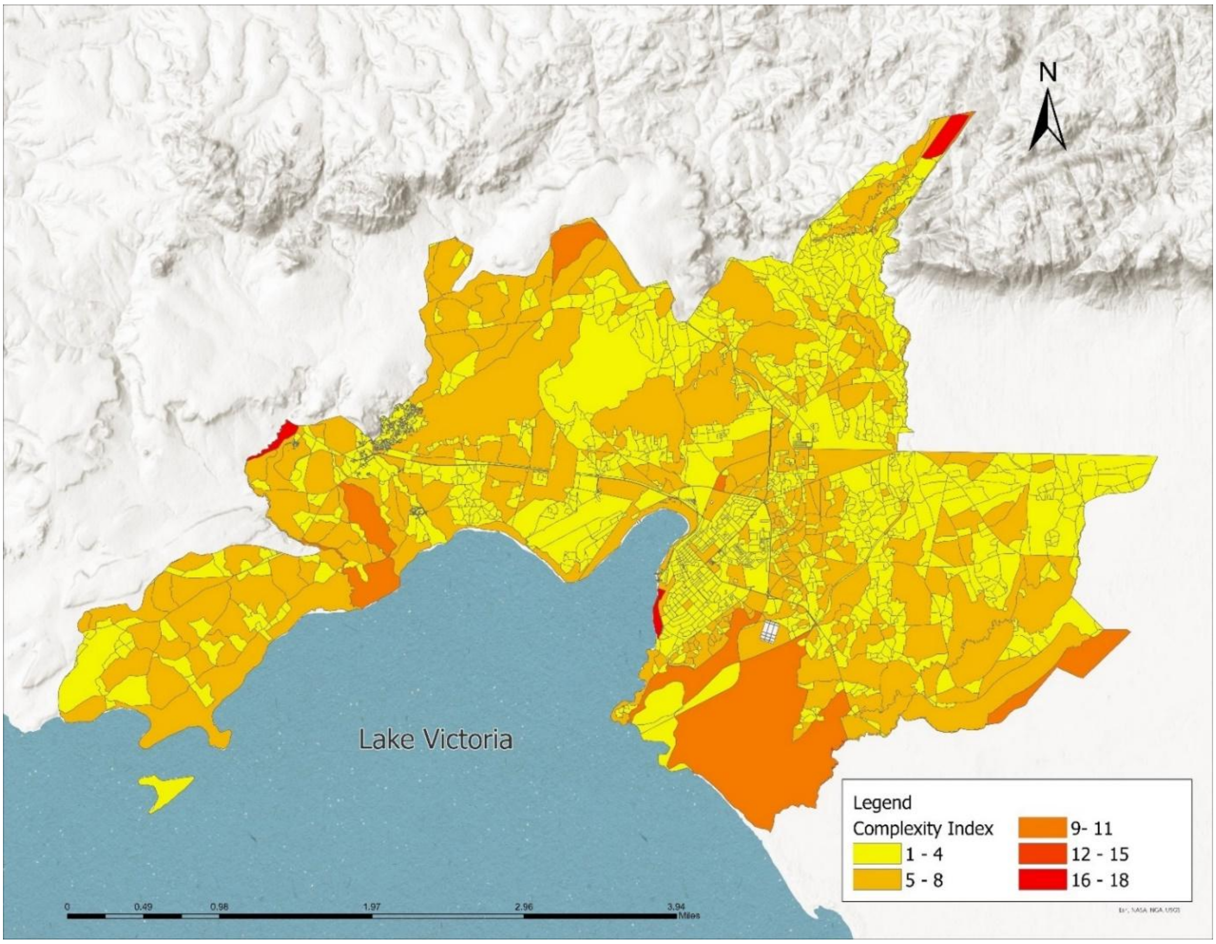


Figure 5.2 The million neighborhoods map of Kisumu, constructed based on the urban footprint of 2020. A higher value for a city block indicates the higher probability of that block being a slum. Data sources: Million Neighborhoods map, 2023 version (Brelsford et al., 2018).

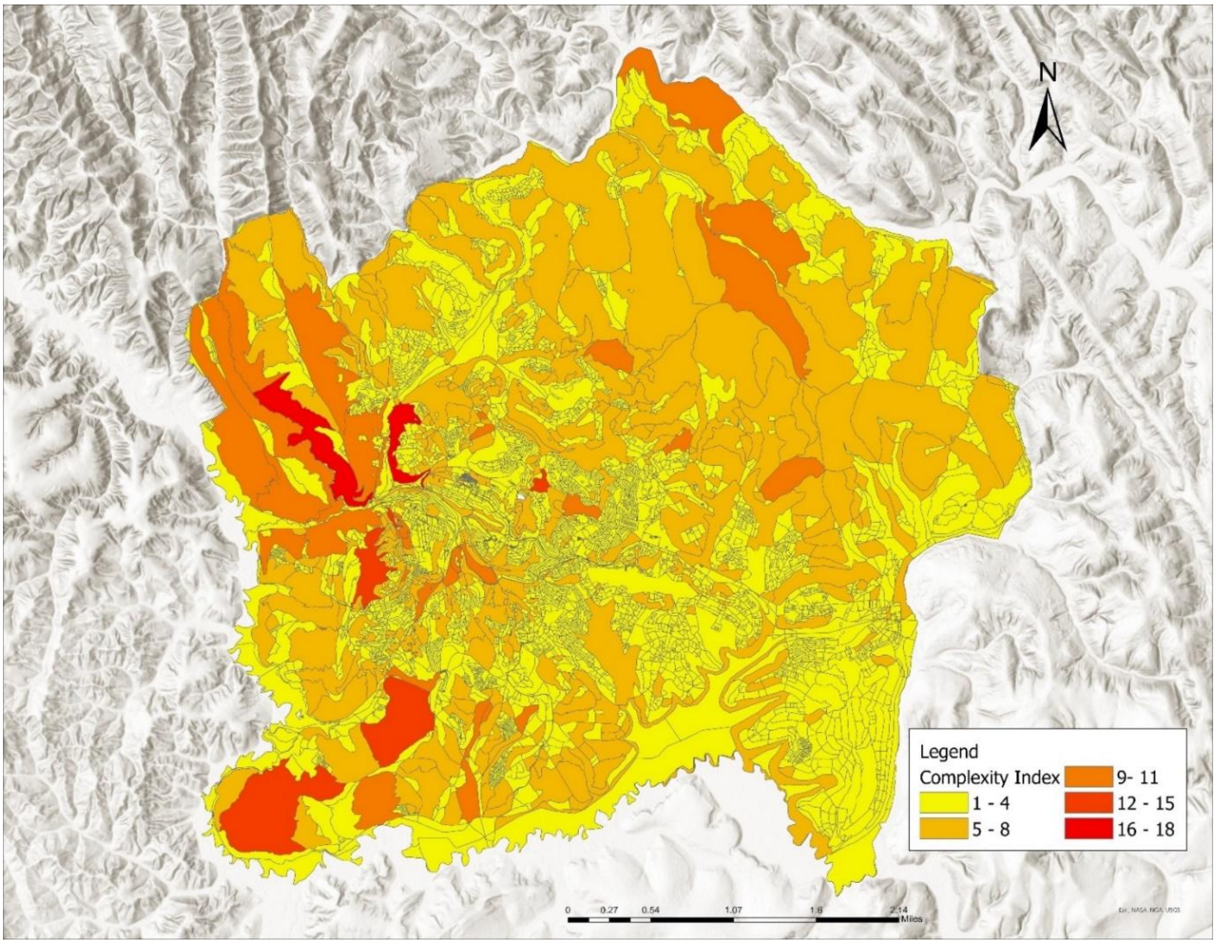


Figure 5.3 The million neighborhoods map of Kigali, constructed based on the urban footprint of 2020. A higher value for a city block indicates the higher probability of that block being a slum. Data sources: Million Neighborhoods map, 2023 version (Brelsford et al., 2018).

To further assess the impact of water management measures on the pipeline network, pipelines were classified based on their management model. Among the two study areas, DMM is exclusively implemented in Kisumu, where it is referenced in the attributes under the 'remarks' field. Kisumu pipelines with the following remarks are managed under DMM: “DMM”, “Managed by DMM”, “Mauna DMM”, “Nyawita Residence DMM Network”, “Pamoja Trust Funded”, “Obukase DMM Network”. These pipelines were delineated for the community similarity analysis. Figure 5.4 shows that a considerable portion of pipelines in slum and rural areas is currently managed under DMM.

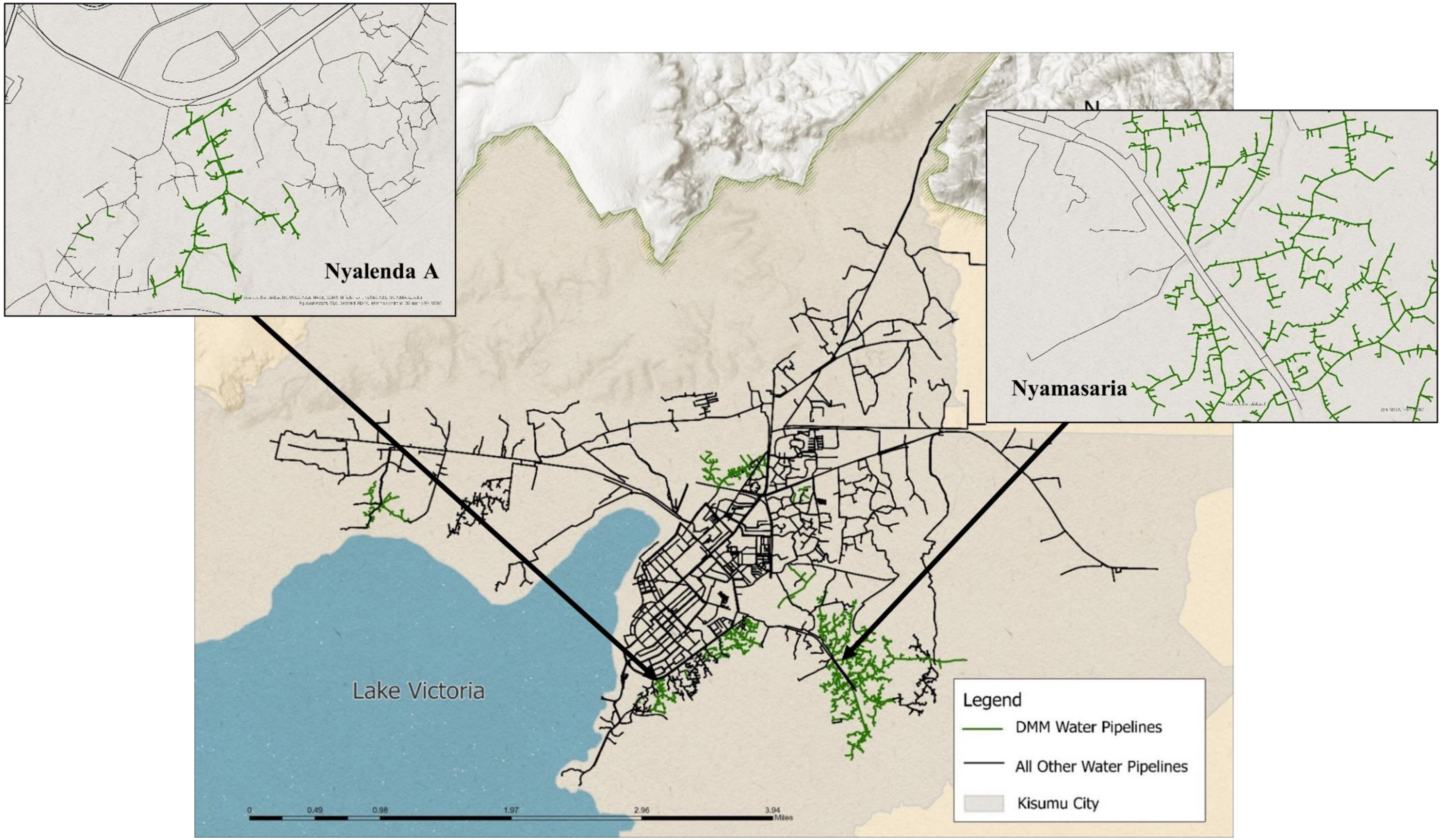


Figure 5.4 Pipelines managed under DMM in Kisumu City, highlighted in green. The inset maps display DMM pipelines located within two neighbourhoods.

5.3.2 Assignment and imputation of weights for Community Detection

Weights can be regarded as additional information about the network. Incorporating weights into graph analysis can yield more accurate assessments of the network, whether in simulating water flow (Scibetta et al., 2013) or evaluating the vulnerability of water pipelines (Delaunois et al., 2014, Kleiner and Rajani, 2001, Mavin, 1996, Wilson, 2010). Furthermore, the use of weights in community detection provides additional insights (Bramson, 2022, Haghbayan et al., 2021, Law et al., 2019). By incorporating weights into community detection, nodes with similar weights tend to cluster in the same community. Considering SDGs, safely managed water services should be available when needed and thus free from the interruptions arising from the breakage of vulnerable pipelines. Therefore, in community detection, the vulnerability of the pipeline material was used to identify communities with similar levels of vulnerability (either structurally or materially) (Figure 5.1).

In a water network, the weights can be physical properties such as geographic distance and pipe diameter or abstract properties such as nodal demand or construction cost (Yazdani and Jeffrey, 2012b). However, the required information for commonly used vulnerability models (Almheiri et al., 2020) cannot be found in the attribute tables of Kisumu and Kigali's pipelines. The attribute fields of Kisumu's pipeline network include diameter, length, roughness, minor loss (i.e., energy loss due to water flow through pipeline components), service status, pipe material, zone, date mapped, installation date, and network type (distribution mains or transmission mains). Additionally, the 'remarks' field captures operational status, operator information, and associated risks. The attributes of Kigali's pipeline network are similar to those of Kisumu but also include the names of pipeline operators. From the statistical analysis of the attributes of the corrected and modelled pipeline networks, it is evident that key attributes related to pipeline conditions, such as roughness, minor loss, and nominal pressure, are either null or contain questionable values across both networks. Statistics on the attributes that are relatively well-preserved in both networks are presented in Table 5.1.

Table 5.1 Statistics on selected properties of case study water pipelines following topological correction

	Kisumu	Kigali
Number of pipelines	4,293	14,172
Average length (m)	123.5842	120.6493
Average diameter (mm)	60.9457	101.34189
Material (percentage of total number)	PVC (84.25%)	PVC (71.81%)
	GI (6.80%)	GI (8.19%)
	Steel (3.61%)	Steel (1.40%)
	PPR (3.07%)	PPR (0.08%)
	AC (1.96%)	Iron (1.55%)
	HDPE (0.30%)	HDPE (11.56%)
		DI (3.51%)
		Missing data (1.91%)
Date mapped/year of installation (years/percentage of total number)	2014 (24.64%)	1890-2022
	2015 (14.98%)	
	2016 (15.70%)	
	2017 (10.34%)	
	2018 (0.63%)	
	Missing data (33.71%)	Missing data (47.40%)
DMM (percentage of total number)	39.00%	N/A

**Note: Missing diameter values are not included in the statistical calculations.*

Based on the statistical results, the material and diameter of the pipelines are the two attributes with the highest completeness. This study, therefore, adopted the index table proposed by Marzouk et al. (2015). The weight index can be used for the evaluation of the pipeline vulnerability even in the absence of detailed data. They estimated the factors that play a significant role in the pipeline's deterioration based on feedback from water experts and calculated the weights for different factors. The weight index for the study pipeline networks was obtained, as shown in Table 5.2.

Table 5.2 Factor weights and grading scales for creating network weights reflecting pipeline vulnerability to breakages (derived from Marzouk et al. (2015))

Factor	Factor weight	Factor grade scales (Scores)				
		1	2	3	4	5
Pipe material type	0.05	PVC	Steel	Ductile Iron	Cast Iron	Asbestos
Pipe diamete (mm)	0.09	>300	300	200	150	≤ 100

Table 5.2 presents two key factors for assessing the condition of pipelines: material and diameter, with vulnerability weights of 0.05 and 0.09, respectively, among a total of 20 factors. The grade scale corresponds to the contribution of different materials and diameters to pipeline deterioration, where 1 represents minimal contribution, and 5 represents the maximum. Based on the provided index, the weights for the pipelines in Kisumu and Kigali can be calculated.

The pipelines in both cities are mainly made of plastic pipelines, i.e. PVC (Polyvinyl Chloride), HDPE (High-density Polyethylene), and PPR (Polypropylene Random Copolymer) pipelines. The materials used in the Kisumu network are PVC, Steel, GI (Galvanized Iron), AC (Asbestos-cement), PPR, HDPE (Table 5.1). PVC, PPR, and HDPE were classified as types of plastic, whereas GI, despite its name, is often made from mild steel sheets and was therefore classified as steel. The materials used in Kigali's pipelines are similar to those in Kisumu's network, including DI (Ductile Iron), Iron, GI, PVC, PPR and HDPE (Table 5.1). Table 5.1 shows that the average pipeline lengths in Kisumu and Kigali are similar, but Kisumu has a smaller average diameter. Subsequently, the two grade scales for each pipeline are multiplied by their corresponding factor weights and summed to obtain the vulnerability weight for that pipeline, as suggested by Marzouk et al. (2015).

Although the material and diameter are the two most complete attributes of Kisumu and Kigali pipelines, there are still gaps in the Kigali network, with 138 missing diameter records and 271 missing material records. Due to the small number of pipelines with missing properties, a multiple imputation method, the sequential imputation of missing values (IMPSEQ) (Verboven et al., 2007a) was used to impute both fields. This method is considered the best approach for imputing missing data in the water distribution system (Kabir et al., 2020, Osman et al., 2018).

The basic principle of IMPSEQ, similar to other multiple interpolation methods, is to impute results via iteration. The IMPSEQ method treats the missing values in the pipeline networks as Missing at Random (MAR), a prerequisite for many imputation techniques (Lin and Tsai, 2020, Newman, 2014, Woods et al., 2024). In other words, it estimates missing data values based on observed variables. The IMPSEQ method divides the dataset into missing matrices D_{miss} and complete matrices D_{com}, and sorts the variables among D_{miss} based on the number of missing values. The variables with the least missing values are estimated first by minimising the determinant of the covariance matrix of D_{miss}, where (Verboven et al., 2007b)

$$cov(D_{miss}) = \frac{t-1}{t} cov(D_{com}) + \frac{1}{t}(D_{miss} - \overline{D_{com}})(D_{miss} - \overline{D_{com}})'$$

And *t* is the number of observed values， $\overline{D_{com}}$ is the average of the observed values.

The above steps are repeated until the imputation of the dataset is complete.

5.4 Topological analysis

In this chapter, the evaluation of the pipeline networks in Kisumu and Kigali is primarily based on their graph metrics and community distribution (Figure 5.1). The former focuses on the overall network structure, while the latter examines the organization of nodes within the network. The same graph metrics as in the previous chapter were used, but, due to the focus on urban features, dual mapping was not applied. Further exploration of these results, including geospatial re-mapping, circuit extraction, and graph similarity analysis, can reveal how the pipelines perform and function in the real world.

Due to the lack of flow data, hydraulic records and empirical data, Kisumu and Kigali's networks were treated as unweighted and undirected in graph analysis (see section 4.2.1). The metrics were calculated using the *networkx* (Hagberg et al., 2008) and *igraph* libraries (Csardi and Nepusz, 2006).

5.4.1 Extraction of pipeline circuits

In addition to evaluating the overall performance of the network, some of the graph theory metrics listed in Section 4.3.1 can also be used to assess the importance of individual pipelines within the network. Urban pipeline networks consist of transmission and distribution systems, with the former connecting water sources to the distribution systems and the latter supplying water to individual users. From a topological perspective, the branches of the pipeline network that do not contain loops are called trees (Deuerlein, 2008). In a tree graph, there is only one path between any two nodes, and subgraphs of the tree are also trees, making it possible to identify them through graph generalisation. As mentioned in Section 2.4.1, several generalisation methods exist. Here, the threshold-based approach was applied to generalise the graph by iteratively removing nodes with a degree of 1, along with their connected edges (Zhou et al., 2010), which are also called leaves (Mair et al., 2017). This method has been applied in simplifying both pipeline and road networks (Hwang and Lansey, 2017, Maschler and Savic, 1999, Pung et al., 2022).

Additionally, edge betweenness was used to highlight important edges in the network. The definition of edge betweenness centrality is similar to that of node betweenness centrality—it represents the number of shortest paths passing through a given edge. In infrastructure networks, edges with high betweenness centrality are critical high-capacity routes (Giustolisi et al., 2019, Yamaoka et al., 2021). Therefore, pipelines with high betweenness centrality were highlighted to represent the key topological structure of the network.

5.4.2 Community detection

5.4.2.1 Selection of algorithms

The algorithm selected for community detection should exhibit high accuracy, meaning that the output partitions should closely align with real-world communities. However, algorithms behave differently between networks, so there are no universal guidelines for selecting algorithms (Ghasemian et al., 2019, McCarthy et al., 2019, Peel et al., 2017). In the context of this chapter, where the community structure of the real network is unknown, based on the literature review (Sections 2.4.2 and 2.4.3), the Infomap algorithm was chosen for the following reasons:

1. **Breadth and depth of algorithm use:** InfoMap has been applied to a wide range of application scenarios(Farage et al., 2021, Hong and Yao, 2019, Hu et al., 2021, Mangioni et al., 2020, Velden et al., 2017), generating algorithmic insights, improvements and robust software for its implementation (Smiljanić et al., 2023).
2. **Performance:** Lancichinetti and Fortunato (2009) consider Infomap to be one of the best performing algorithms. Agreste et al. (2017) also point out that “Infomap algorithm showcased the best trade-off between accuracy and computational performance.” The comparison by Wickramasinghe and Muthukumarana (2022) also indicates that Infomap performs well in sparse networks. Given that infrastructure networks are often sparse, this makes the Infomap algorithm a suitable choice for the analysis.
3. **Resolution of output detected communities:** Infomap has the advantage of resolution. Fortunato and Barthélemy (2007) indicate that many modularity-based algorithms suffer from an inherent resolution problem in that they cannot identify communities smaller than a certain size. Their minimum resolvable community size depends on the total size of the network and the degree of interconnection of the modules. This is because modularity-based algorithms (For example, the Newman-Girvan modularity described in the literature review 2.4.2) work by comparing network clusters with those in a random network. As the network size increases, the expected number of connections between clusters in the random network decreases. Therefore, when connections exist in the

clusters of a large-scale network, they will be treated as strong connections, resulting in two communities with different characteristics being merged. In contrast, Kawamoto and Rosvall (2015) demonstrate that Infomap performs well in terms of resolution.

4. **Hierarchical community classification:** Real-world networks are often hierarchical, and Infomap can detect their hierarchical structure. In large systems, there is often a hierarchical relationship between communities, with smaller communities being part of communities at a higher level. This structure can improve the efficiency and stability of the system. Therefore, identifying the hierarchical structure is essential to understanding the network's performance. Many other algorithms do not have this capability (Lancichinetti et al., 2011).

This study used the multi-level method of Infomap. If a community is still subdividable, according to the map equation, the multi-level Infomap algorithm will continue to partition the network until all communities are indivisible, thus dividing the network into various levels of communities. Due to Infomap being a heuristic algorithm, the trial number was set to 50 to obtain a best-performing partition (with the shortest description length) from the 50 calculations. In Infomap, Markov time refers to the coding frequency of the random walker's path, thus controlling the expected community size. The Markov time was set to 1, as Poorthuis (2018) suggests that a Markov time in the range of 0.6 to 1.1 helps to detect meaningful city structure. Setting it to 1 means that the random walker's position is encoded at each step. In addition, unweighted graphs were used to evaluate the impact of incorporating vulnerability index values as weights into the InfoMap workflow.

5.4.2.2 Validation of the community's structure

Infomap outputs communities at different levels, and only some are meaningful in a real-world context. Therefore, without prior knowledge of true community structure, it is important to find reliable methods to evaluate the output of the community detection algorithm (Signorelli and Cutillo, 2022). The Community Structure Validation (CSV) index proposed by Signorelli and Cutillo (2022) was used in this study to test the structural strength of the output communities. This approach follows similar principles to the modularity and Order Statistics Local Optimization Methods (OSLOM) (Lancichinetti et al., 2011) in assuming that the density of connections within a community is higher than the density of connections between communities. It compares the observed connections with a hypergeometric null model, as proposed by Lancichinetti et al. (2010). For undirected networks, the hypergeometric null model is defined as (Signorelli et al., 2016):

$$N_B \sim hypergeom(n = d_A, K = d_B, N = d_V)$$

Where V is the node set of a graph, whilst A and B are node subsets A and B. d_A, d_B, d_V denote the total degrees of sets A, B and V.

CSV evaluates two types of hypotheses for each community pair C_r and C_s:

Internal density of a community C_r is tested by: $H_0: \mu_{rr} = \mu_{rr}^0$ vs. $H_1: \mu_{rr} > \mu_{rr}^0$

where μ_{rr} denotes the expected number of links between nodes in C_r, and μ_{rr}^0 is the corresponding null expectation from the hypergeometric model.

External sparsity between communities C_r and C_s is tested by: $H_0: \mu_{rs} = \mu_{rs}^0$ vs. $H_1: \mu_{rs} < \mu_{rs}^0$

where μ_{rs} denotes the expected number of links between nodes in sets C_r and C_s, and μ_{rs}^0 is the corresponding null expectation from the hypergeometric model.

If both null hypotheses are rejected in large proportions for a given type I error α, it upholds the idea that there is a clear community structure in the network. In other words, given the significance level, the CSV index specifies to what extent the structure of the network conforms to the definition of a community based on statistics describing the connections in a network. The CSV ranges from 0 to 1, and the closer the value is to 1, the clearer the community structure. In addition, Signorelli and Cutillo (2022) introduced the weighted CSV, which weights the CSV based on the strength of rejecting the null hypothesis. Like the original CSV, it ranges from 0 to 1. Both CSVs were applied in this study. The CSVs are influenced by the size of the network, and given the size of the Kisumu and Kigali networks, their CSVs results should be reliable according to Signorelli and Cutillo (2022)'s tests. The code is available from https://github.com/mirkosignorelli/csv.

5.4.2.3 Partition similarity

The interpretation of the output of community detection can be seen as an extension of Section 5.4.2.2. In this context, the DMM attribute and the million neighbourhoods map (complexity index) serve as ground truth partitions. Measuring the distance between these partitions can further uncover the similarity between network communities and the real-world environment (Bramson, 2022, Law et al., 2019). According to Fortunato and Hric (2016), techniques for measuring the distance between two partitions can be categorised into pair counting, cluster matching and information theory. Commonly used metrics include the fraction of correctly detected vertices, adjusted Rand index (ARI), adjusted mutual information (AMI), and normalised mutual information (NMI) (Danon et al., 2005, Girvan and Newman, 2002, Liu et al., 2019, Vinh et al., 2009).

In this chapter, the metrics used to measure the distance between partitions are ARI and AMI (Vinh et al., 2009). The Rand index measures the overlap of communities to indicate the similarity between two partitions, while mutual information assesses the entropy of two discrete variables (partitions) (Liu et al., 2019). In other words, the mutual information measures the information needed to infer a partition given the other partition (Dao et al., 2020). The term 'adjusted' refers to the adjustment for chance. In the test of the Rand index and mutual information, Vinh et al. (2009) found that an increase in the number of communities in a partition leads to higher Rand index and mutual information values that are independent of ground truth. ARI and AMI correct this bias, ensuring that the ARI and AMI of randomly generated partitions are fixed given ground truth. In other words, the accuracy of the two metrics is not affected by the number of communities. Both metrics range from 0 to 1, with 1 meaning that the two partitions are identical and 0 meaning that the similarity between the two is the same as that of any two random graphs.

In this chapter, the blocks in the Million Neighborhoods map, whose boundaries are from Maxar Technologies Inc and Ecopia.AI, were taken as communities. These blocks were dissolved based on the complexity index classification of Brelsford et al. (2018), and the IDs of the dissolved blocks were given to the pipeline network nodes that spatially overlapped with them to form a new partition. The partitions were further refined in the analysis of the impact of DMM on the pipeline network, where nodes within the same dissolved community that belong and do not belong to the DMM pipeline are further partitioned into different communities. In other words, partitions related to slum distribution were generated based on the Million Neighborhoods Map, and the Kisumu partition was refined according to DMM information by creating additional communities. Following division, the average degree, betweenness, and closeness centrality of pipeline nodes within areas of varying block complexity and within DMM areas were computed as an initial comparison of communities. Then, both sets of partitions were then used to measure the relevance of pipeline topology to the distribution of slums and DMM areas using ARI and AMI.

5.5 Results

In the results section, the global performance of the water networks was first presented by evaluating graph metrics of the Kisumu and Kigali primal graphs. The subsequent section focuses on the centrality measures derived from these metrics and interprets their distribution in relation to the characteristics of each city. As one of the most widely recognised graph metrics in network analysis, centrality has been demonstrated to provide insights into the configuration of key pipelines that are critical for transmission. The differences in centrality performance between Kisumu and Kigali were explored. The final parts analyse the community structure of the study area networks. Communities, representing tightly connected groups within the network, often correspond to real-world regional characteristics. Therefore, this section examines the relationship between slum distribution and network communities within the study areas and further interprets the connection between Kisumu's pipeline communities and the DMM.

5.5.1 Graph analysis for Kisumu versus Kigali

The pipeline networks of Kisumu and Kigali were simplified through graph generalisation and pipelines with high edge betweenness centrality were highlighted in Figures 5.5 and 5.6. The simplified pipeline networks in both cities form circuits surrounding vital urban areas. In Kisumu, the simplified pipelines are predominantly located in the southern part of the city (Figure 5.5), whereas in Kigali, the density of major pipelines decreases from west to east (Figure 5.6). The overlap of high centrality pipelines with the simplified network suggests that circuit pipelines play a crucial transmission role and thus have higher importance. Furthermore, the maps show differences in the pipeline layout within slums/unplanned settlements between Kisumu and Kigali. Kisumu slums have more tree-like structures that are excluded from the simplified map, while Kigali has a more connected network in slums.

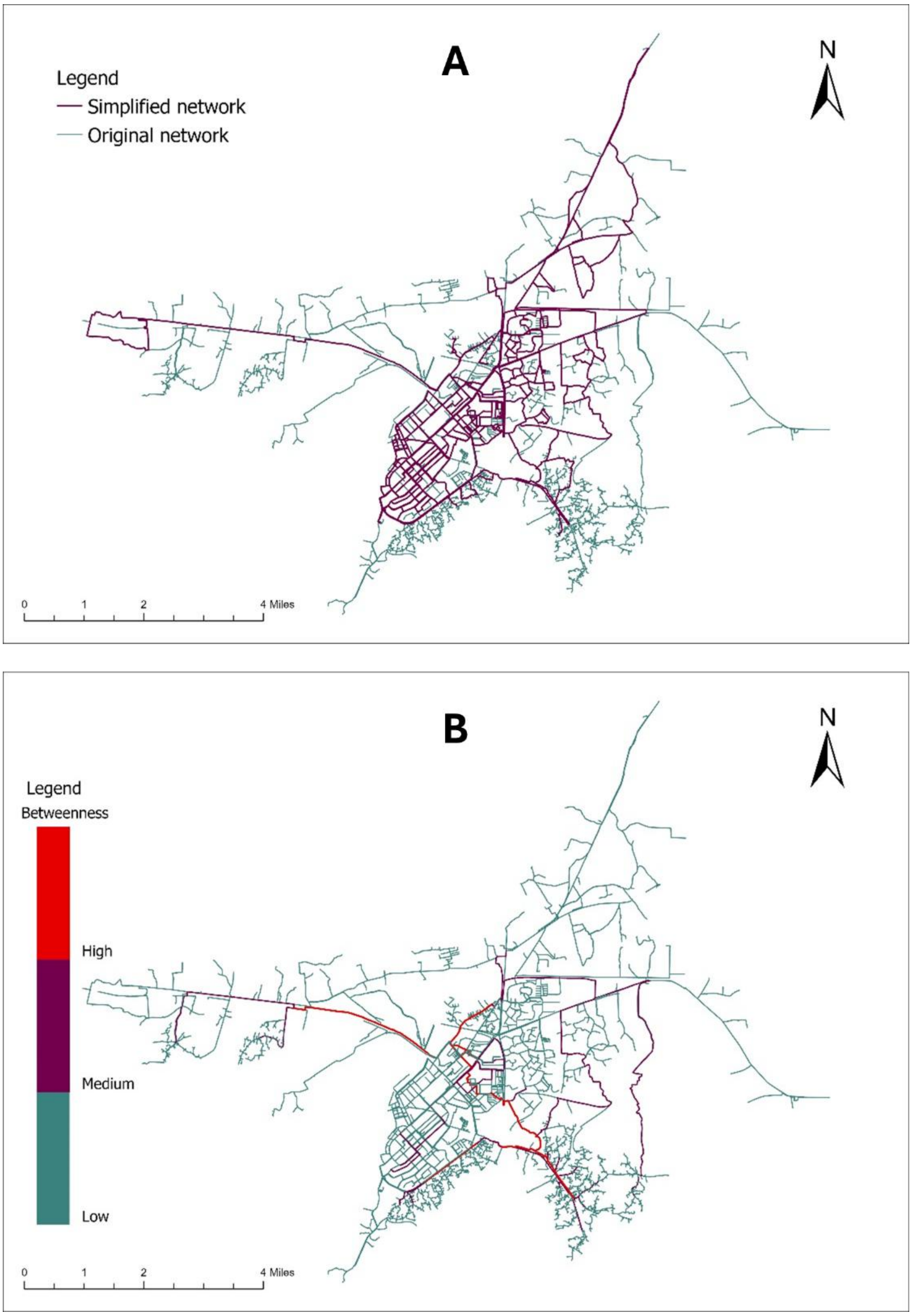


Figure 5.5 Maps of simplified pipeline (A) and edge betweenness centrality distribution (B) in Kisumu. Both distributions highlight critical pipelines from the perspectives of degree and betweenness.

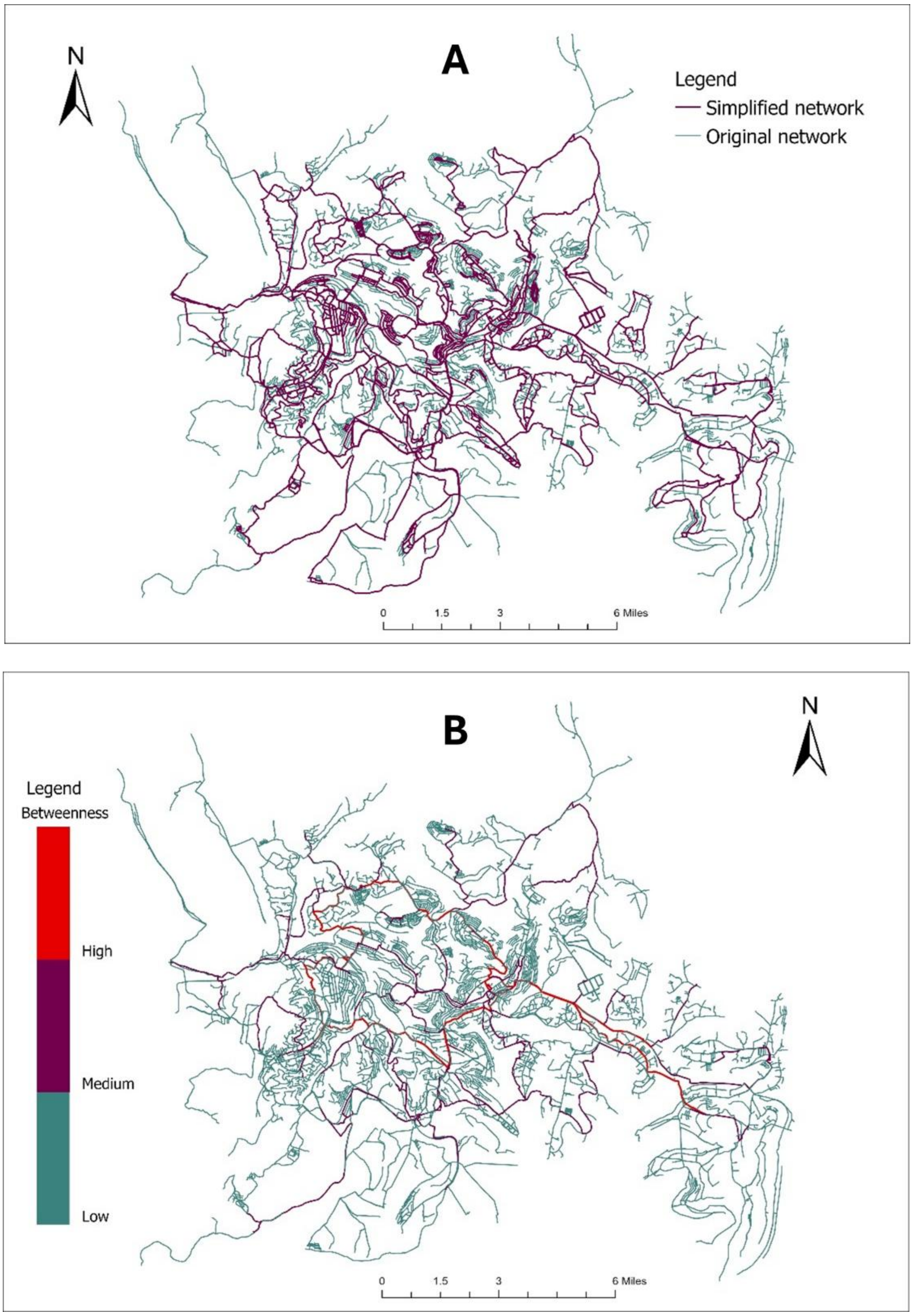


Figure 5.6 Maps of simplified pipeline (A) and edge betweenness centrality distribution (B) in Kigali. Both distributions highlight critical pipelines from the perspectives of degree and betweenness.

Table 5.3 provides additional key properties of the pipeline networks. Firstly, it can be seen from the number of nodes and edges that after smoothing and fixing the topology, Kigali has a larger network size than Kisumu. Nevertheless, despite differences in network scale, most graph metrics between the two cities show only minor differences, demonstrating the robustness of the graph metrics under various city contexts.

Table 5.3 Graph attributes and measurements for the water networks

	Kisumu	Kigali
Nodes	3,911	6,254
Edges	4,076	6,721
Average node degree	2.0844	2.1493
Normalized betweenness centrality	0.0100	0.0060
Normalized closeness centrality	0.0264	0.0267
Cluster coefficient	0.0054	0.0143
Meshedness coefficient	0.0212	0.0374
Average shortest path length	40.0610	38.7395
Algebraic connectivity	0.0001	0.0001
Spectral gap	0.2715	0.0010

The average node degree hardly varies with the network size. According to Giudicianni et al. (2018), it ranges from 2 to 4.5 in water distribution networks, with a lower value representing less robustness. The table shows that both have relatively low average node degrees, suggesting the low redundancy characteristic of water systems, as Wéber et al. (2020) demonstrated, and Kisumu's network is more sparsely connected (Yu et al., 2024).

The average shortest path length and clustering coefficient are used to assess the orderliness of a network. The average shortest path length measures the steps required to link nodes across the network. At the same time, the clustering coefficient indicates the number of loops or faces of a network, reflecting how tightly nodes are connected. Ordered networks typically score high on both metrics. Kigali's clustering coefficient and meshedness are considerably higher than Kisumu's. Combined with the slightly larger degree and smaller average shortest path length, metrics suggest that the Kigali network is connected more efficiently. This aligns with Kigali's development emphasis on planning for urban infrastructure (see section 3.2). Additionally, the clustering coefficient and average shortest path are indicators of small-world networks, which are widely observed in real-world systems and characterized by dense local clusters connected by few inter-cluster links (Schnettler, 2009). Kigali's higher clustering coefficient and shorter average path length suggest that its network is more likely to exhibit small worldliness compared to Kisumu. However, specific metrics are required to formally identify a small-world network (Neal, 2018), so this conclusion remains tentative.

Spectral gap and algebraic connectivity are low in both networks, which is expected due to the inherent sparsity of infrastructure systems. Spectral gap indicates how the network is connected. The larger the network, the smaller spectral gap generally is, as small spectral gap values suggest more critical bottlenecks or bridges that can split the network into two or more isolated parts (Yazdani and Jeffrey, 2012a). Based on this metric, it can be concluded that Kigali lacks long pipelines connecting distant regions, to which network size and terrain may contribute. However, Kigali surprisingly has greater algebraic connectivity than Kisumu, meaning network bisection is more challenging in Kigali (Zeng et al., 2017). It also implies that there are no clear high-density and low-density clusters in Kigali, or that its pipeline network has more evenly distributed clusters, as evidenced by the low average shortest path length.

5.5.1.1 Network centrality

Degree, betweenness, and closeness are three commonly used network metrics for assessing edge importance, making it essential to measure them further. Although it is not possible to use the distribution to describe the pattern of degrees due to the small range, it can be noted that both networks tend to be right-skewed. As illustrated in Figure 5.7, nodes with a degree of 2 in both networks have a low frequency, primarily due to the removal of pseudo-nodes during graph smoothing. Additionally, Figure 5.7 indicates that degree distributions in the Kisumu and Kigali networks exhibit a high frequency of nodes with degrees 1 and 3, which correspond to endpoints and the T-junction of pipelines respectively. In other words, the networks may have many tree-like branch pipelines.

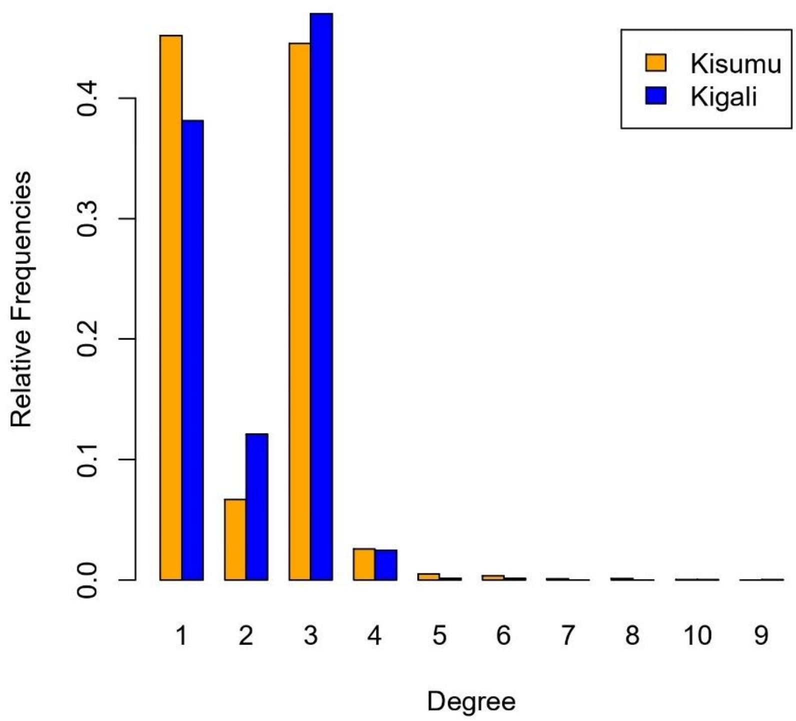


Figure 5.7 Degree distribution of the Kisumu and Kigali networks.

Figures 5.8 and 5.9 show similar distributions for betweenness and closeness centrality, which are well-represented by Weibull and lognormal distributions. This finding is consistent with the reported centrality distribution of roads in Hong Kong (Lan et al., 2022). However, the functions of the two networks have subtle differences. The Weibull function better fits Kisumu's betweenness and closeness with low Akaike information criterion (AIC) and Bayes information criterion (BIC). In contrast, the lognormal distribution performs better for Kigali's centrality data. This suggests that, for both networks, the betweenness and closeness are heavy-tailed, and there are "backbone" pipelines with high centralities in the networks (Fornito et al., 2016a).

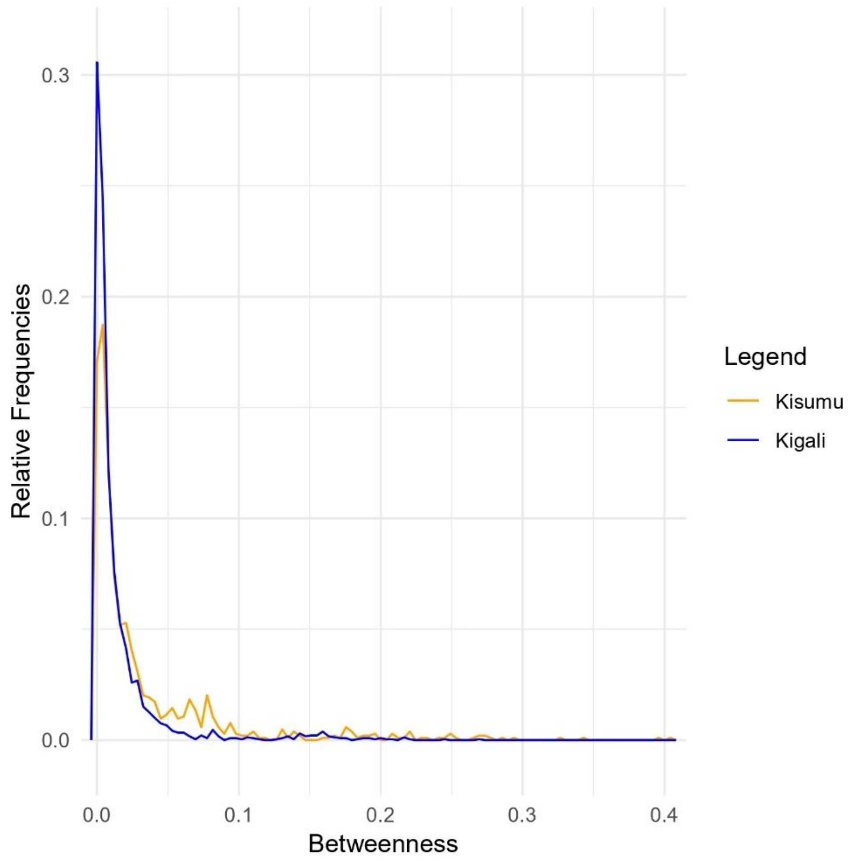


Figure 5.8 Betweenness distribution of the Kisumu and Kigali networks.

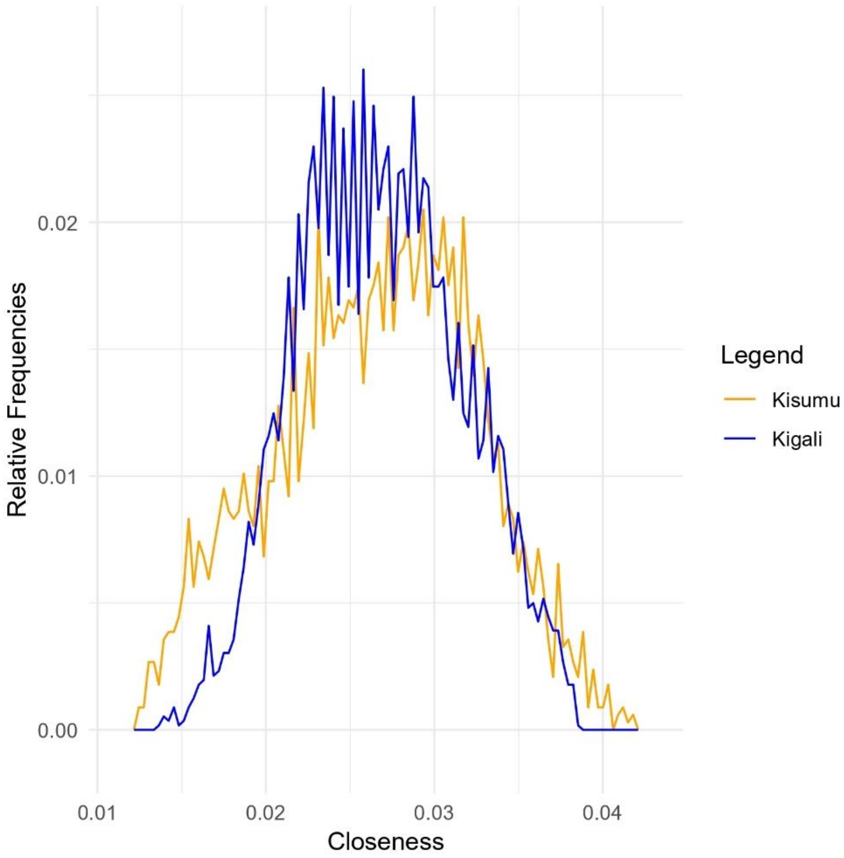
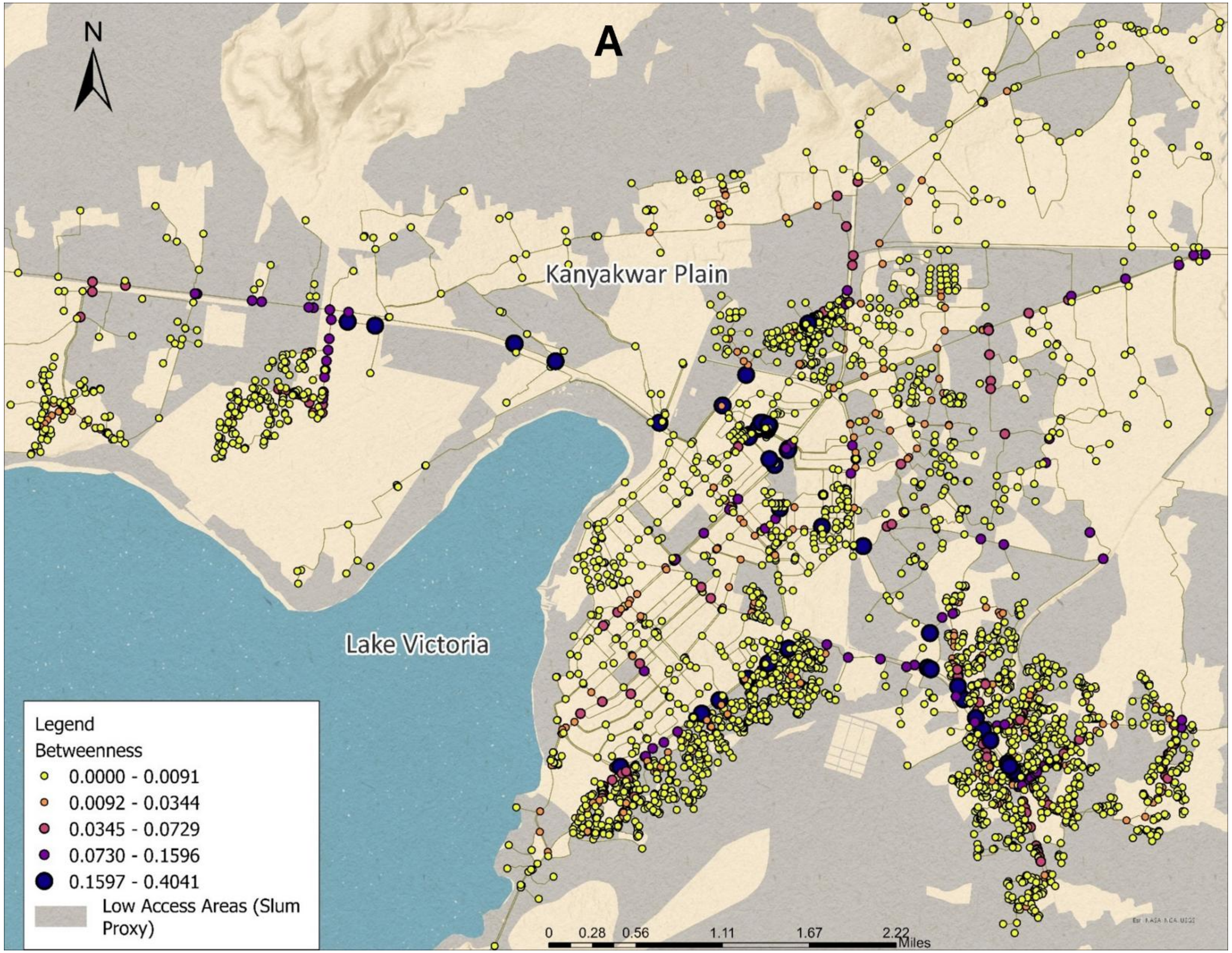


Figure 5.9 Closeness distribution of the Kisumu and Kigali networks.

The centrality maps (Figures 5.10 and 5.11) show the difference between betweenness and closeness in the two study areas more clearly. In both Kigali and Kisumu, nodes with high betweenness centrality are the minority and are distributed regularly, the same as the distribution shown in Figure 5.8. In Kisumu, high betweenness nodes are distributed along the direction of its urban sprawl (County of Kisumu, 2013). If compared with the Million Neighborhoods map, it can be observed that these points are also distributed at the boundaries of informal and formal areas, where pipelines and buildings form distinct neighbourhood boundaries. In Kigali, the high betweenness pipelines roughly form a Y-shape, which overlaps considerably with the three main pipelines outlined in the Kigali master plan (City of Kigali, 2020a). Nodes with high closeness centrality that can effectively connect to others are mainly located in the central area of Kisumu. In particular, a gradual decrease in closeness can also be found at the boundary between Kigali's formal and informal areas, with low closeness nodes roughly distributed within the unplanned settlements.


A
N
Kanyakwar Plain
Lake Victoria
Legend
Betweenness
0.0000 - 0.0091
0.0092 - 0.0344
0.0345 - 0.0729
0.0730 - 0.1596
0.1597 - 0.4041
Low Access Areas (Slum Proxy)
0 0.28 0.56 1.11 1.67 2.22 Miles

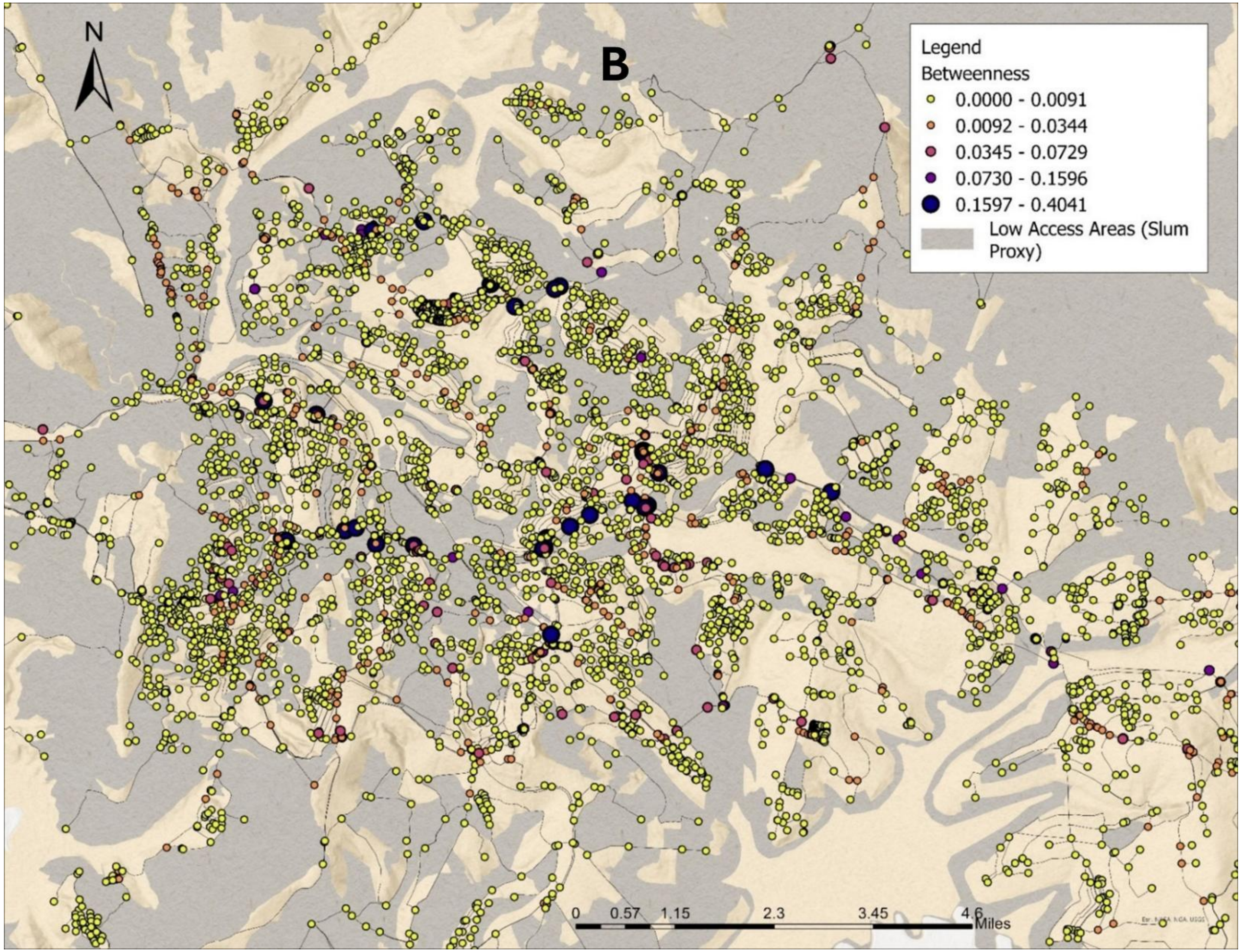
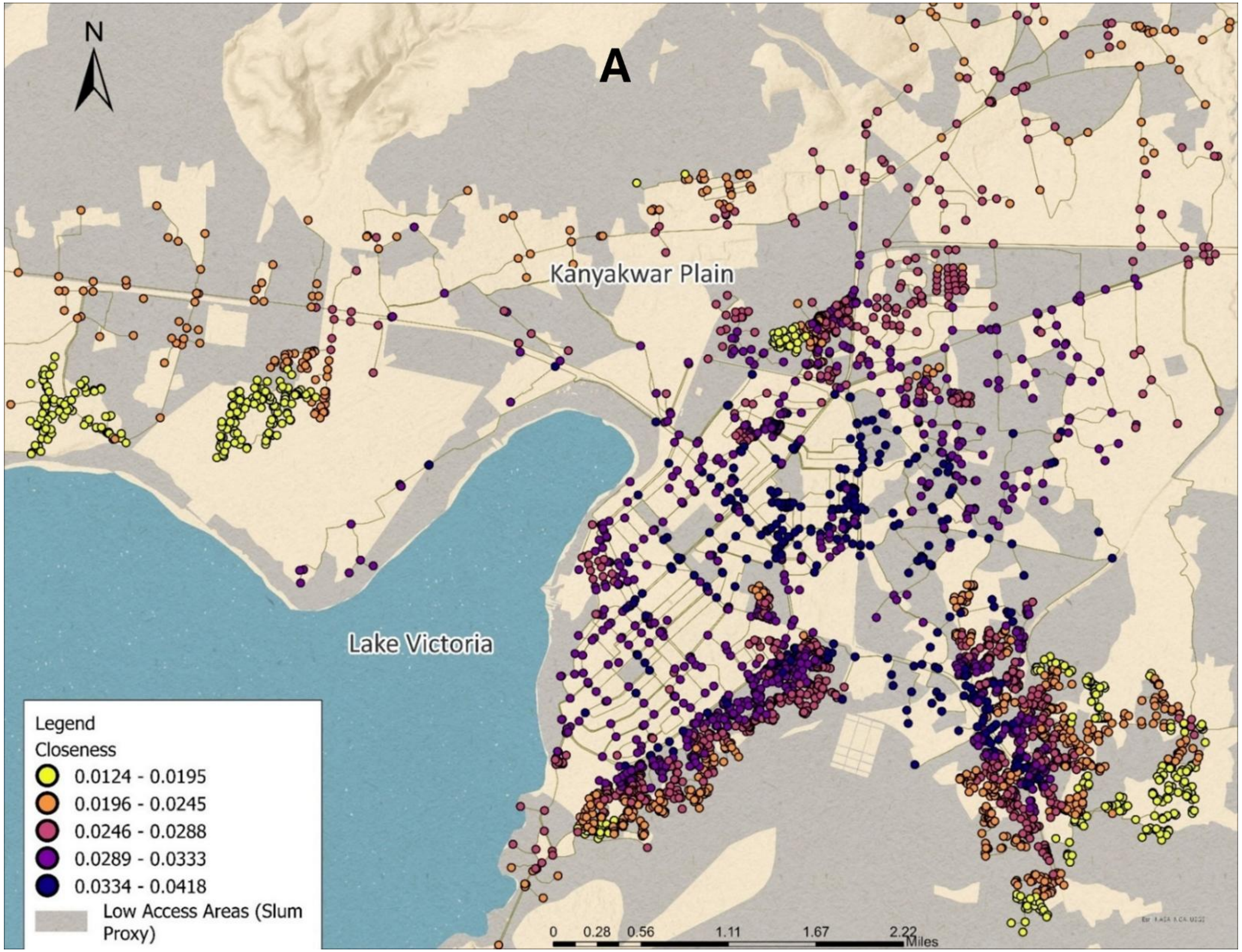


Figure 5.10 Betweenness centrality maps for Kisumu (A) and Kigali (B). Betweenness centrality indicates the importance of a pipeline in connecting various parts of the network. Darker colours represent higher centrality values.


N
A
Kanyakwar Plain
Lake Victoria
Legend
Closeness
0.0124 - 0.0195
0.0196 - 0.0245
0.0246 - 0.0288
0.0289 - 0.0333
0.0334 - 0.0418
Low Access Areas (Slum Proxy)
0
0.28
0.56
1.11
1.67
2.22
Miles

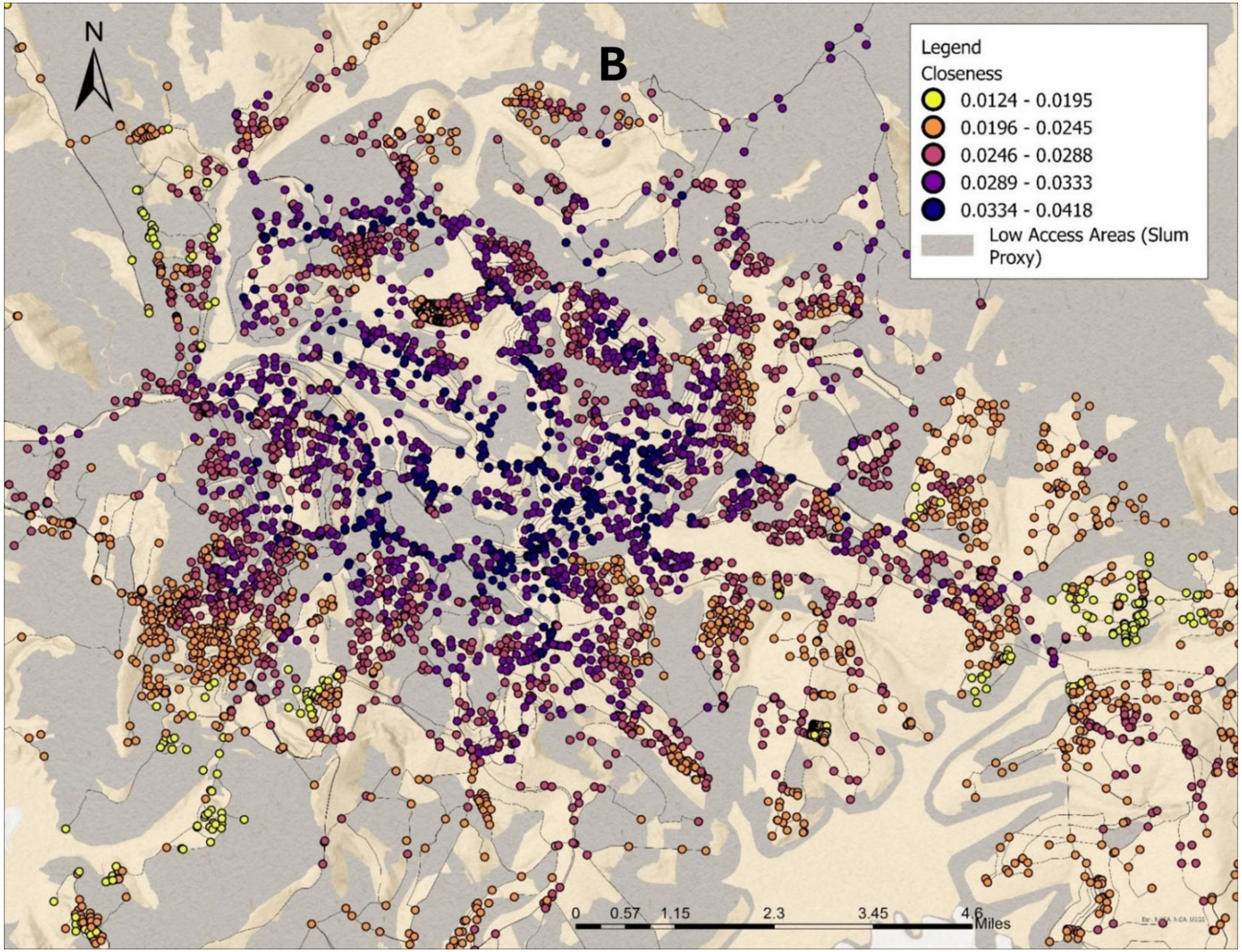


Figure 5.11 Closeness centrality maps for Kisumu (A) and Kigali (B).Closeness centrality reflects the ease with which a pipeline can reach all other nodes. Darker colours represent higher centrality values.

5.5.2 Validation and interpretation of the results of community detection

Table 5.4 shows that the community detection stopped at the seventh layer for Kisumu's weighted network, the eighth layer for Kisumu's unweighted network, and the eighth layer for both Kigali's weighted and unweighted networks. Both weighted and unweighted CSVs were employed to assess the outputs at these layers. The CSV values range from 0 to 1, with higher values indicating a partition that is more likely to exhibit a clear community structure. Since the CSV calculation excludes clusters with fewer than 5 nodes, the values fluctuate as the levels increases and the granularity of the output becomes finer. However, considering the scale of the study network, the initial few layers of Infomap output, which are identified as having strong community structure based on the CSVs, remain reliable. According to Table 5.4, Infomap identifies community structures more efficiently in Kisumu's weighted network as it reaches the maximum CSV values faster. It can be seen that the CSV values reach the maximum of the Kisumu networks in the weighted first level (Figure 5.12) and unweighted second level, with these levels having a similar number of communities. Kigali is the opposite, as its unweighted rather than weighted network reaches the maximum CSV values more quickly (Figure 5.13). Moreover, Kigali and Kisumu have similar numbers of communities at the layers where their maximum CSV values occur.

Table 5.4 Number of communities detected from case study water piped networks using the Infomap algorithm, together with associated CSV index values.

Hierarchical level		Kisumu			Kigali		
		Number of communities	CSV	Weighted CSV	Number of communities	CSV	Weighted CSV
Weighted network	Level 1	14	1.000	1.000	2	0.667	0.667
	Level 2	71	0.466	0.358	15	1.000	1.000
	Level 3	328	0.023	0.016	92	0.378	0.302
	Level 4	1365	0.009	0.009	479	0.013	0.009
	Level 5	3096	0.037	0.036	2182	0.005	0.005
	Level 6	3805	0.400	0.400	4957	0.015	0.015
	Level 7	3911	n/c	n/c	6170	0.250	0.250
	Level 8	n/c	n/a	n/a	6254	n/c	n/c
Unweighted network	Level 1	2	0.667	0.667	9	1.000	1.000
	Level 2	12	1.000	1.000	52	0.654	0.573
	Level 3	74	0.361	0.270	306	0.046	0.029
	Level 4	378	0.016	0.011	1526	0.005	0.004
	Level 5	1766	0.010	0.009	4346	0.010	0.010
	Level 6	3151	0.023	0.023	5927	0.065	0.063
	Level 7	3879	0.400	0.400	6223	0.500	0.500
	Level 8	3911	n/c	n/c	6254	n/c	n/c

Figure 5.12 Kisumu Thiessen polygon map of weighted communities at level 1, where the CSVs reach their maximum value.

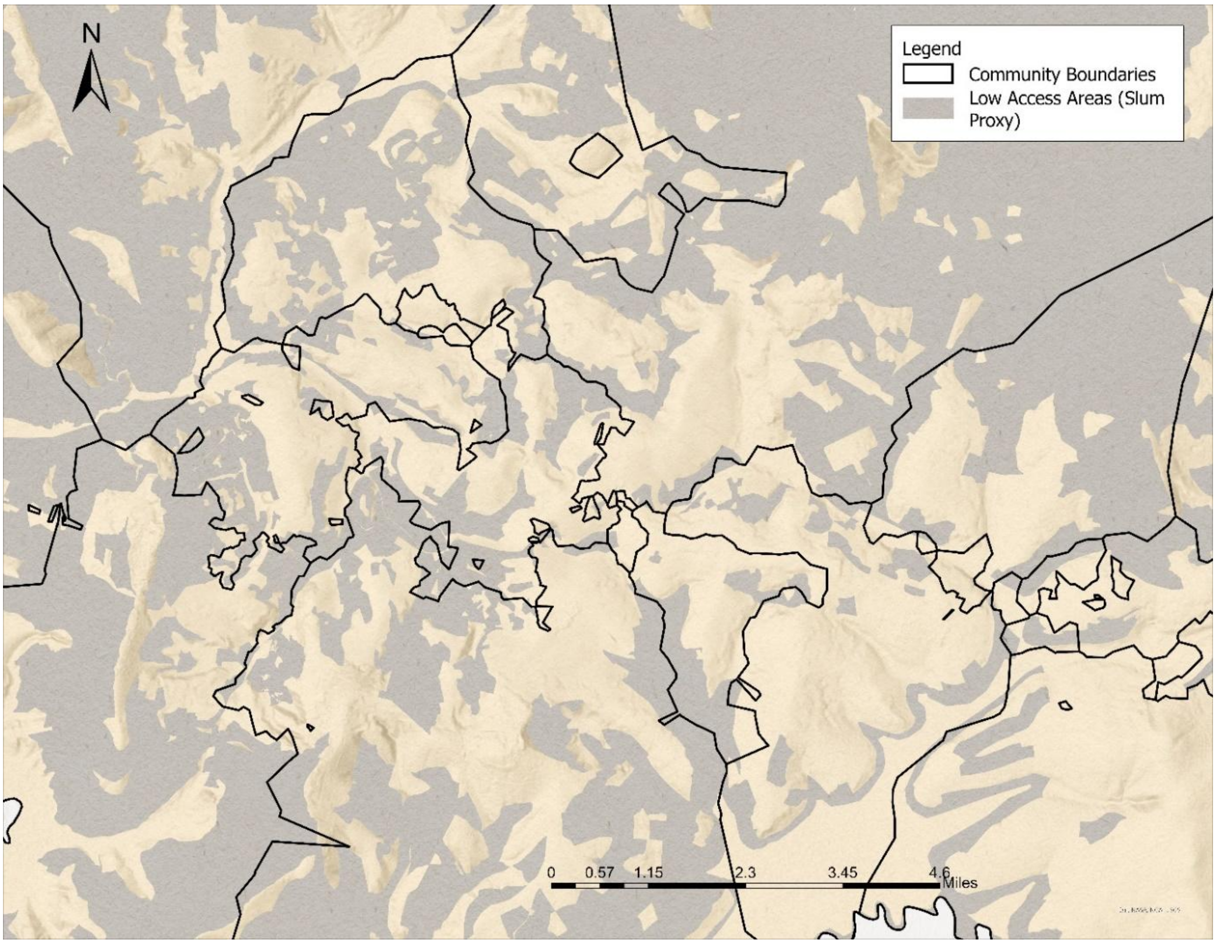


Figure 5.13 Kigali Thiessen polygon map of weighted communities at level 2, where the CSVs reach their maximum value.

5.5.2.1 Kisumu

As shown in Table 5.5, the number of pipelines in Kisumu initially increases and then decreases with rising levels of neighbourhood morphological complexity. This trend contrasts with the distribution of the number of neighbourhoods across the same complexity levels, which exhibits a consistent decline—from 2,287 blocks at complexity level 1, to 910, 233, and 52 blocks at levels 2, 3, and 4, respectively. Although exact density cannot be computed due to variations in neighbourhood size, it is evident that most pipelines are located within moderately complex areas.

When viewed through the lens of DMM, the total number of pipelines under DMM and non-DMM management in Kisumu is roughly equal. However, the proportion of DMM-managed pipelines increases with morphological complexity. The degree and betweenness of DMM pipeline nodes are generally lower than those for other areas with the same settlement complexity level. This indicates that DMM pipelines primarily serve as branch lines for water distribution to consumers. In addition, both average degree and betweenness centrality decline as settlement morphological complexity increases and areas become more slum-like, with the decline being particularly pronounced for betweenness. On the other hand, closeness centrality remains relatively stable across all areas, suggesting that it is less associated with local settlement structural characteristics.

Table 5.5 Statistics of centrality metrics for Kisumu pipelines, categorised by neighbourhood morphological complexity and DMM status

Complexity Class	DMM*	Count	Mean Degree	Mean Betweenness	Mean Closeness
1	0	250	2.3080	0.0144	0.0292
	1	54	2.2037	0.0234	0.0212
2	0	1007	2.1887	0.0132	0.0268
	1	708	2.0508	0.0085	0.0256
3	0	582	2.0928	0.0109	0.0264
	1	905	1.9746	0.0064	0.0255
4	0	169	1.9704	0.0099	0.0274
	1	236	1.9576	0.0048	0.0289

**DMM code: 0 indicates non-DMM pipelines; 1 indicates DMM-managed pipelines.*

ARI and AMI are two metrics used to measure the similarity between partitions. Figure 5.14 show that the trends in ARI and AMI for the Kisumu partitions are similar across levels. Overall, the value of AMI is always higher than the ARI for the same partition, and both AMI and ARI are higher when DMM partition is used as the "ground truth" compared to when only slum data is used. The figure also shows some subtle trend variations. The AMI of the weighted graph peaks at community level 2 (Figure 5.15), while the unweighted graph shows a rightward trend, peaking at community level 3, which has a similar community number to the second layer of the weighted graph. These findings suggest that weighting the pipeline network aids in the discovery of real-world communities. The largest ARI values of slums and DMM are both observed in the second layer of the weighted network. Interestingly, the ARI peaks for slums and DMM areas from the unweighted network are observed in the second layer rather than the third.

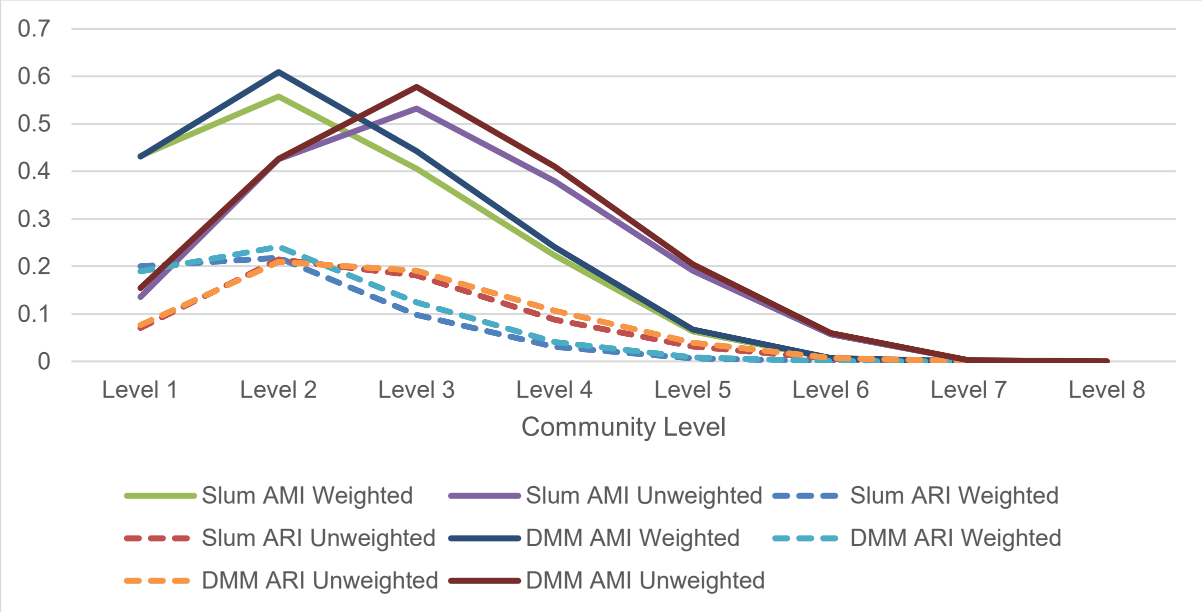


Figure 5.14 AMI and ARI distributions for weighted and unweighted communities in Kisumu. The figure illustrates the distance between Kisumu's pipeline communities and partitions derived from the Million Neighborhoods map, as well as the DMM partitions refined from the former.

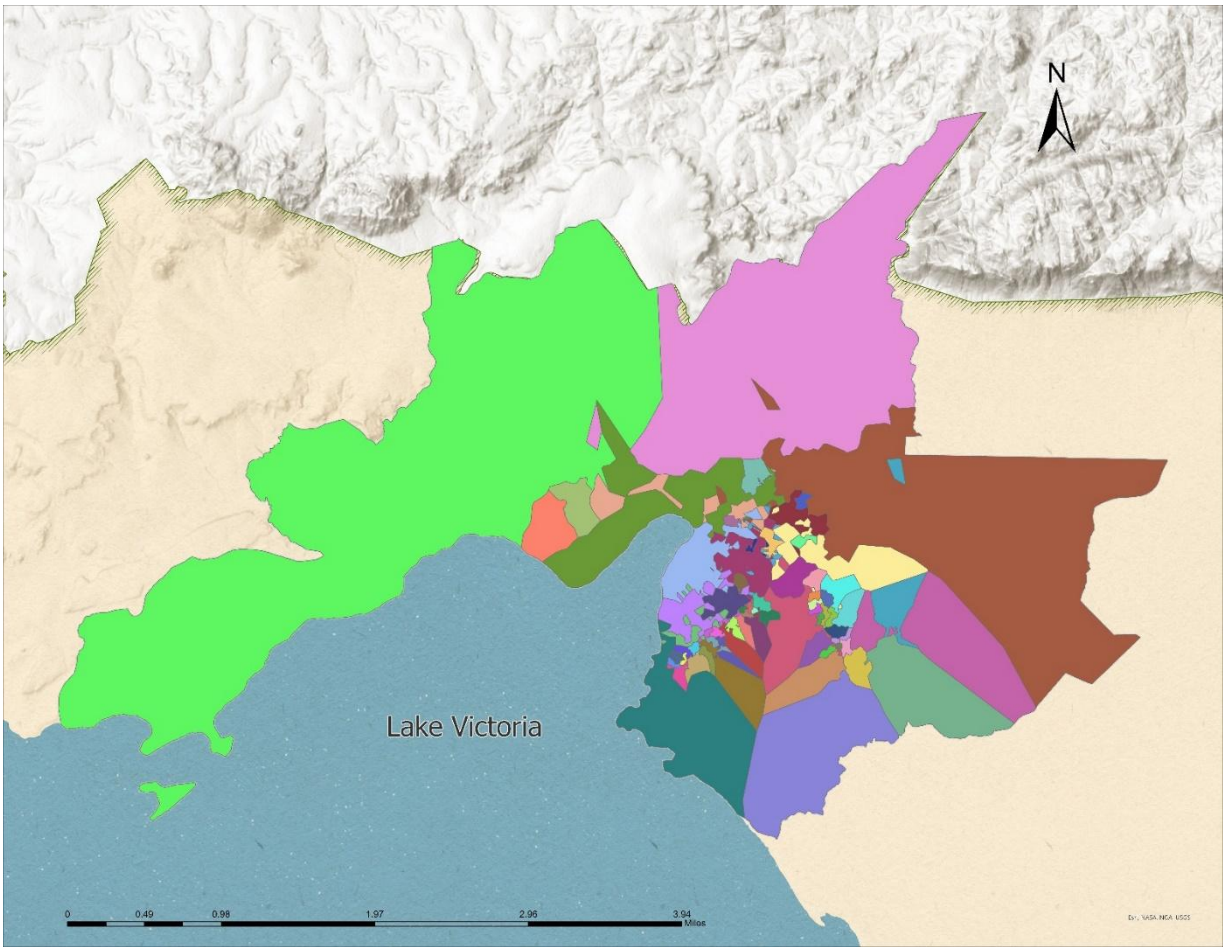


Figure 5.15 Kisumu Thiessen polygon map of weighted communities at level 2, where the AMI and ARI reach their maximum values.

5.5.2.2 Kigali

Table 5.6 reveals a markedly different pattern of pipeline distribution in Kigali compared to Kisumu. On the one hand, the number of neighbourhoods in Kigali also decreases with increasing morphological complexity, from 4,536 at level 1 to 2,137, 376, and 119 at levels 2, 3, and 4, respectively. However, in contrast to Kisumu, pipelines in Kigali are densely concentrated in the least complex neighbourhoods. In addition, the centrality metrics of pipelines in Kigali exhibit the same trends observed in Kisumu: both degree and betweenness centrality decline with increasing morphological complexity, while closeness centrality remains relatively stable.

Table 5.6 Statistics of centrality metrics for Kigali pipelines, categorised by neighbourhood morphological complexity.

Complexity Class	Count	Mean Degree	Mean Betweenness	Mean Closeness
1	2192	2.2359	0.0069	0.0267
2	2603	2.1425	0.0058	0.0267
3	842	2.0689	0.0062	0.0268
4	617	1.9806	0.0038	0.0266

Figure 5.16 shows that Kigali also has a higher AMI than ARI. The key difference between Kigali and Kisumu lies in how AMI and ARI respond to the weighting and unweighting of the networks. In contrast to Kisumu, Kigali's weighted networks are more right-skewed than unweighted networks with AMI/ARI peaking at level 3 for weighted networks (Figure 5.17) but level 2 for unweighted networks, a pattern also reflected in statistics. In addition, there is no clear relationship between the maximum values of AMI and ARI. The maximum ARI occurs in the third level of the weighted network and the second level of the unweighted network, corresponding to the CSV of communities. The third level of both graphs has the maximum AMI.

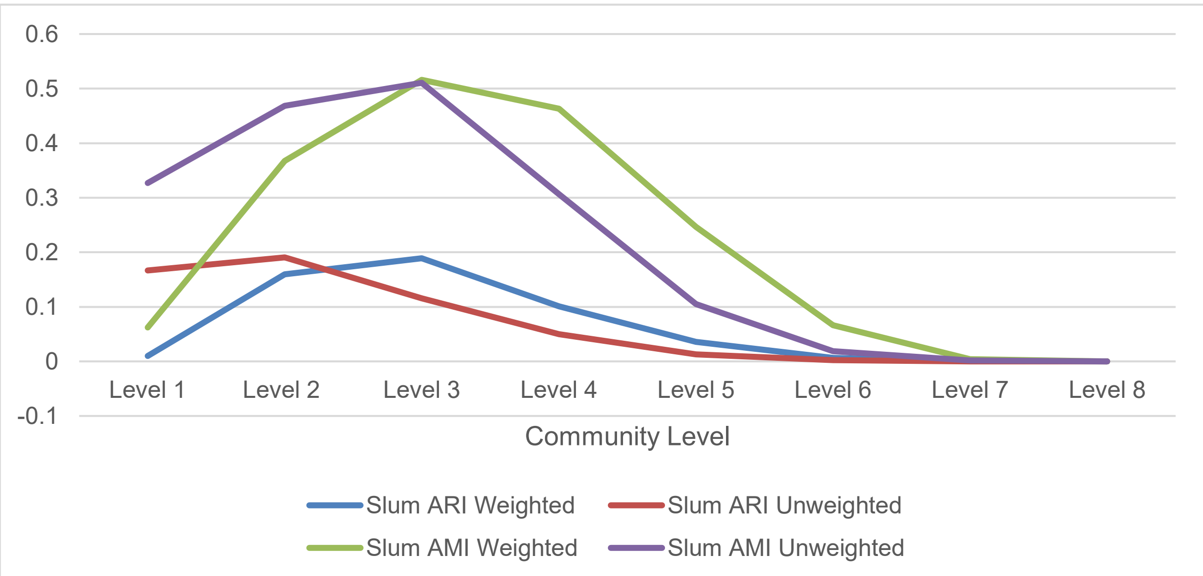


Figure 5.16 AMI and ARI distributions for weighted and unweighted communities in Kigali. The figure illustrates the distance between Kigali's pipeline communities and partitions derived from the Million Neighborhoods map.

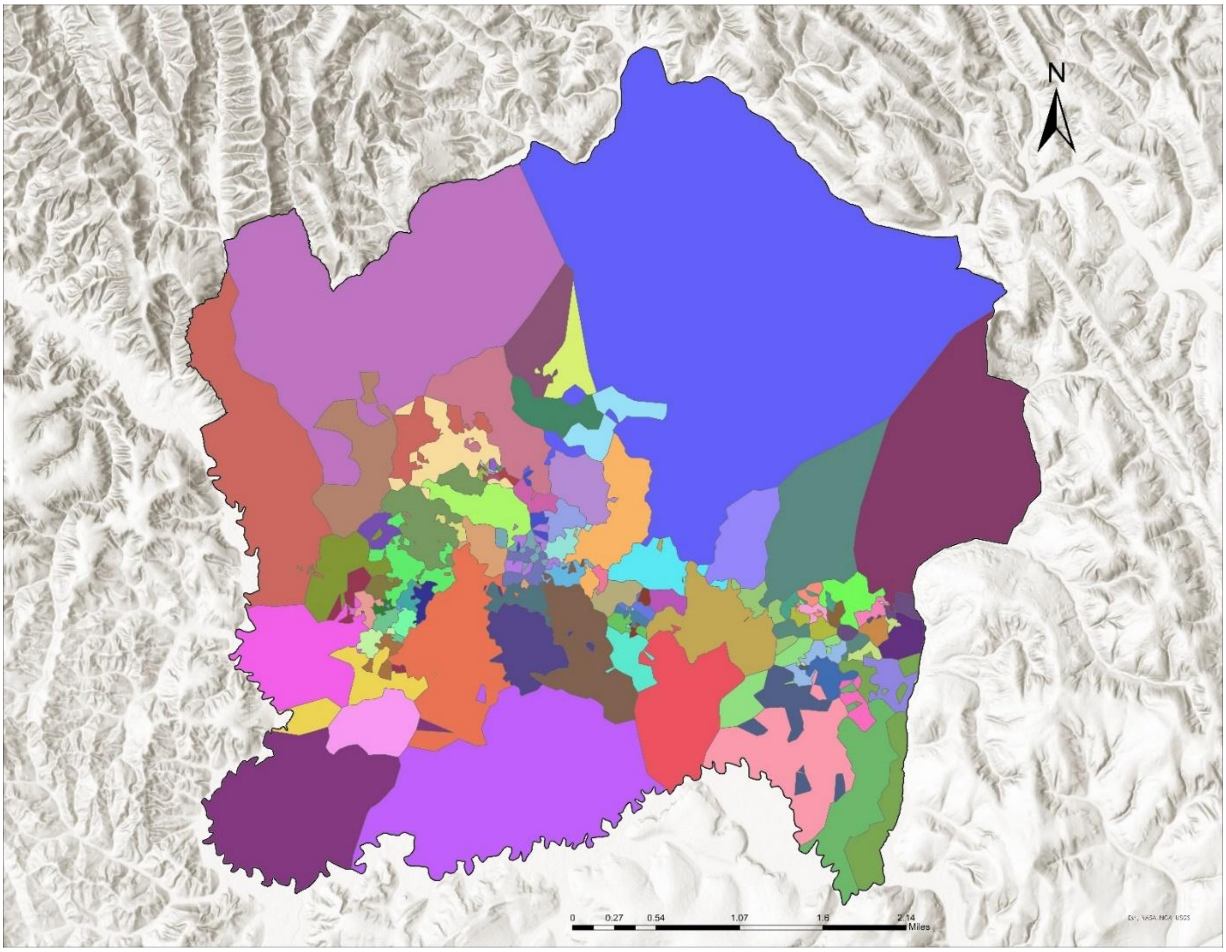


Figure 5.17 Kigali Thiessen polygon map of weighted communities at level 3, where the AMI and ARI reach their maximum values.

5.6 Discussion

In this study, graph theory and community detection methods were applied to the water pipeline networks of Kisumu and Kigali to investigate their performance, hierarchy, community structure, and vulnerability. By converting the primary elements of the pipeline networks into graph nodes and edges, the relationship between the networks and the urban environment was highlighted. Compared to other water pipeline networks studied in literature (Giudicianni et al., 2018), both cities have lower average node degrees (close to 2), indicating line graph properties with relatively lower robustness. Other metrics from Table 5.3 show that Kisumu has a more uneven pipeline distribution than Kigali. Furthermore, Kigali's network is more similar to the small-world network, with dense clusters and fewer connections between them.

In the graph simplification, more branch pipelines were excluded in Kisumu's slums than in Kigali's. Since the excluded pipelines have tree structures that are less costly but less reliable in infrastructure networks (Han et al., 2020), the simplification of the pipeline networks implies differences in planning tendencies within the slums of the two cities. Kigali's well-developed water system in slums aligns with the city's overall development plan. Uwizeye et al. (2022), Benken (2017), Jaganyi et al. (2018) and Hudani (2020) have found that Kigali has implemented urban plans incorporating various policies, such as the National Urbanisation Policy (MININFRA, 2015b) and Vision 2020 (GoR, 2012). These initiatives aim to transform Kigali into an intrinsic economic growth driver while fostering inclusivity in the post-genocide period (Manirakiza et al., 2019). As a result, these policies have objectively facilitated access to public services, including water, sanitation, electricity, and land registration, for residents of Kigali's slums. Simplified pipeline maps and centrality distribution maps also show significant spatial structural differences between formally planned settlements and slums in both cities. High betweenness centrality nodes are concentrated along the main urban arteries and the boundaries of the city's central areas, overlapping with urban loops. The closeness centrality of both cities' networks decreases from their centres outward to the slums.

Statistical results for centrality emphasize the similarities between the two study networks. Both cities have a high number of intersections with a degree of 3 (Figure 5.7), consistent with the degree distribution patterns observed in road networks of over eighty global cities (Badhrudeen et al., 2022). The betweenness and closeness centrality distributions of the two networks are similar and can be represented via Weibull and lognormal distributions, which are flexible and can be transformed into Gaussian and power-law distributions. This supports Broido and Clauset (2019)'s conclusion that many networks are better described by distributions other than the power-law one. However, there are also topological differences between the study networks. On the one hand, as network size increases, the connectivity of water distribution networks tends to decrease, making them more prone to fragmentation (Giudicianni et al., 2018). On the other hand, urban infrastructure networks are influenced by urban morphology and historical factors (Nor et al., 2021, Yazdani and Jeffrey, 2012a). For example, Xue et al. (2022) and Mengistie et al. (2023) observe that cities in developed countries have significantly more homogeneous road networks than those in developing countries. Similarly, Kut et al. (2016) note that cities with similar historical backgrounds may exhibit similar graph-theoretic properties. Kisumu's development has always been centred around the Kanyakwar Plain near Lake Victoria, expanding outward from its colonial city core. Kisumu's role as a transportation hub linking eastern and western Kenyan cities has significantly influenced its urban growth and infrastructure development (County of Kisumu, 2013, County of Kisumu, 2020, LVSWSB, 2021) , which explains the spatial configuration of its pipelines and why its high-betweenness pipelines do not align with large-diameter north-south pipelines (Figure 5.5). Meanwhile, although Kigali also experienced social and spatial heterogeneity, urban planning has played a significant role in its urbanization (Ananga et al., 2019, Baffoe et al., 2020a, Baffoe et al., 2020b, LVSWSB, 2021, Manirakiza et al., 2019). This may account for the subtle variations in algebraic connectivity and centrality distributions between the two cities.

The distribution of centrality metrics across the two cities is not without pattern. As shown in Tables 5.5 and 5.6, the properties of pipelines are closely associated with the morphological characteristics of neighbourhoods. In Kigali, pipelines are concentrated in the least complex neighbourhoods; the better planned a neighbourhood is, the more pipelines it tends to contain. This pattern differs from that observed in Kisumu, reflecting the divergent urban planning strategies of the two cities. The trend in Kigali corresponds to a city-wide strategy that emphasises the upgrading of entire neighbourhoods into higher-standard residential areas, rather than implementing targeted interventions in water infrastructure alone. Conversely, the pipeline distribution in Kisumu suggests a pattern of infrastructural intervention in response to specific settlement conditions. Notably, in both cities, the degree and betweenness centrality of pipelines decrease with increasing morphological complexity, regardless of the overarching planning strategies. This supports the conclusion drawn in Chapter 4—namely, that the spatial properties of urban networks are closely tied to neighbourhood typologies. It is also worth highlighting the role of DMM-managed pipelines, whose proportion increases with neighbourhood complexity. These pipelines tend to exhibit lower betweenness centrality, indicating that they primarily function as branch lines within informal settlements.

Community detection sheds further light on the relationship between urban environments and water infrastructure. A key conclusion is that the distribution and vulnerability of pipelines in Kisumu and Kigali are closely related to the distribution of slums. From a graph theory perspective, the pipeline networks in Kisumu and Kigali can be partitioned into a similar number of weighted communities (Table 5.4). Given that Kigali's network is larger than that of Kisumu, the fact that both cities have a similar number of detected communities supports the view that Kigali's pipeline system is more evenly distributed. Moreover, graph similarity analysis, considering slums as "ground truth", shows that the topology and vulnerability of the pipelines in both cities are strongly related to slum distribution. This result agrees with Buhl et al. (2006), who observed that street networks in slum areas exhibit unique characteristics. The AMI and ARI values (Figures 5.14 and 5.16) indicate that Kisumu has more heterogeneous pipeline networks in slums than Kigali. Notably, higher AMI and ARI values are concentrated in the first three partition levels, after which both measures gradually decline. This trend is reinforced by the CSV results, which indicate that partition structures are strongest in these initial layers, supporting the interpretation that early partitions capture meaningful features of the network relevant to slum distribution.

Additionally, graph similarity findings suggest that Kisumu's DMM policies are also associated with the layout of pipelines. DMM pipelines operate as distinct subsystems managed and constructed by master-operators, aiming to replace the chaotic 'spaghetti' network with a more structured one (World Bank, 2009). Comparison of pipeline-derived communities with the DMM partitioning of urban space yielded higher similarity scores, indicating that the topology of DMM pipelines distinctly differs from that of other slum pipelines. This finding partially addresses Nzengya (2015)'s concern that there is a lack of evidence to show whether DMM genuinely improves the layout of pipeline networks within slum areas.

These findings contribute to bridging urban science and network studies. Researchers in city science emphasize the need to understand cities and urban planning both through urban processes and infrastructural form, particularly graph theory properties (Brelsford and Martin, 2021). Southern cities face challenges with fragmented networked services, and there is consensus on the impact of urban governance difficulties and slums on these networks. Nevertheless, infrastructure in southern cities, particularly in slums, is often overlooked in graph theory research, and there is limited understanding of the factors influencing urban network layout and how to measure these impacts, especially in developing countries (Neal et al., 2021). This chapter contributes to addressing these issues by providing insights into the pipeline topology of the study areas and developing a workflow for adapting graph theory methods to the characteristics of available data in LMIC cities.

This chapter adopted the same network preprocessing methods and graph metrics as those employed in Chapter 4, but placed greater emphasis on the analysis of the performance of water networks. As a result, the graph approach in this chapter is subject to certain limitations:

1. Graph metrics provide only an initial answer to questions regarding water networks. Both empirical and hydraulic data are still necessary to establish a direct link between network topology and management activities. As Yazdani and Jeffrey (2011) point out, topological analysis alone provides an incomplete picture of network resilience, as financial and operational management, geographical context and the urban space heavily influence the structure of the network.
2. Graph analyses of infrastructure often vary in their modelling methods (Marshall et al., 2018, Pueyo et al., 2019, Zhou et al., 2010), such as whether to use primary or dual mapping, and how to correct and simplify network topology. Each method has its own particular applications, but in practice, this methodological diversity can affect the reliability and generalizability of the results (Giustolisi et al., 2019, Marshall et al., 2018).

3. Many existing studies of water network topology rely on synthetic networks, which do not always reflect the characteristics of real-world networks (Momeni et al., 2023, Paez and Filion, 2017, Yu et al., 2024). As a result, conclusions drawn from graph studies of pipeline networks still require validation through further research based on real networks in SSA regions.

The analytical workflow is also constrained by data limitations. The absence of installation date records for pipeline data restricted this analysis to a cross-sectional study, limiting further exploration of temporal causal relationships between urban environmental change, policies, and subsequent pipeline network evolution. Additionally, inconsistencies in digital pipeline mapping considerably increased the analytical workload, diminishing reproducibility and preventing hydraulic simulations. Therefore, it is encouraged that utilities improve the management of geospatial data, particularly by systematically recording attributes such as pipeline installation dates (even if only approximately for the oldest network segments), materials, and diameters. This will enable more detailed evaluations of pipelines and facilitate spatio-temporal analysis of network evolution.

Furthermore, the study acknowledge the inherent limitations of community validation methods. The method relies on prior knowledge to identify environmental factors. In this study, the assessment of slum impacts is based on the Million Neighborhoods Map. However, in practice, there are multiple approaches to identifying slums (Kuffer et al., 2016, McCartney and Krishnamurthy, 2018, Smit et al., 2017). Improving the accuracy of slum infrastructure analysis will thus ultimately depend on further understanding of slum morphology. The study may also overlook other influential factors related to pipeline distribution. In particular, the mountainous terrain of Kigali undoubtedly influences both settlement distribution and the laying of pipelines, presenting a promising direction for future research. Researchers may consider employing the workflow established in this study while incorporating detailed elevation data.

5.7 Conclusion

The expansion of slums is a key feature of urbanization in SSA. Slums are characterized by dense populations, overcrowded buildings, lack of tenure and security, and chaotic management, all of which have hampered efforts to improve water access. This chapter examined the state of water services in the slums of Kisumu and Kigali by applying graph theory and community detection methods, offering insights into pipeline networks and their relationship with the urban environment in developing countries. Graph metrics provide a valuable overview of pipeline networks, especially in locations where information is scarce, while the Infomap algorithm reveals the networks' hierarchical and clustering structure, indicating their response to urban environmental change and policy. The results show that Kisumu and Kigali share similar topological characteristics common to southern cities. Furthermore, both government intervention and the spatial configuration of the slum are significantly associated with the pipeline layout. The findings contribute to a deeper understanding of urban networks in developing countries and offer insights for improving water infrastructure in such contexts.

Chapter 6 Measuring the spatial accessibility of water kiosks in Kisumu

The structure and performance of pipeline networks represent only part of the urban water landscape. According to the WHO/UNICEF Joint Monitoring Programme (JMP), improved water services are further classified into safely managed, basic, and limited water services based on accessibility, availability, and quality (WHO/UNICEF, 2018). Under these standards, in 2022, 2.2 billion people still lacked safely managed drinking water, with 1.8 billion people lacking on-premises drinking water (UN, 2024). Accessibility on-premises is the most common limiting factor for safely managed drinking water services in Sub-Saharan Africa (SSA). In 2022, nearly half (45%) of the 1.2 billion people in SSA relied on water collected outside their homes, a proportion significantly higher than in other regions (WHO/UNICEF, 2023a). This indicates that a large portion of the population relies on public water sources, such as kiosks, public taps, and boreholes, for basic and limited services. For example, Uganda's 2021 census showed that 8% of households used public taps during the dry season, with this number rising to 11.3% in the wet season—both higher than the proportion of households using piped water in dwellings during the same period (UBOS, 2022). Similarly, Kenya's 2019 census reported that 9.9% of households nationally used public taps or standpipes, which was higher in urban areas, reaching 15.6% (KNBS, 2019b). These shared water sources greatly extend water coverage in low-income communities (Post and Ray, 2020). To comprehensively evaluate water access in urban areas, it is essential to assess these diverse water sources in addition to conducting graph theory analyses of water pipeline networks.

Among shared sources, water kiosks—micro-enterprises that sell piped water to households without direct pipeline connections—play a crucial role in supplying water in many African cities. Despite their importance, assessments of kiosks have primarily relied on field surveys and interviews (Adams, 2018b, Falcone et al., 2023, Nel et al., 2023, Opryszko et al., 2013). While such methods provide valuable insights, they fail to account for the influence of urban-scale environmental factors—such as the distribution of infrastructure and population—on water access. Consequently, they fall short in quantifying the geographic accessibility of kiosks, which is crucial for understanding and improving urban water systems. Thus, this chapter aims to quantify geographic access to kiosk water within Kisumu using the two-step floating catchment area (2SFCA) method. To evaluate how spatial representation of population affects accessibility metrics, the method integrates three gridded population map datasets with water kiosk locations to estimate the supply and demand ratio. Additionally, the relationship between kiosk accessibility and road networks is analysed due to their close interconnection.

6.1 Introduction

Shared or communal water sources have a long history in Africa (Nilsson, 2011) and come in various forms, including standpipes, kiosks, water tankers, household resellers, and water vendors (Chitonge, 2014). The proliferation of shared water sources aligns with the emergence of the informal sector and public-private partnerships (PPPs) in urbanization (Section 2.2.3). Urbanization poses two major challenges in water management: a mismatch between water supply and demand, and the unequal distribution of water resources (Dos Santos et al., 2017, Rebelo and Matos, 2022). Access to improved water sources in SSA is influenced by factors such as household characteristics (e.g. income, gender, and education) as well as urban planning and historical policies (Antunes and Martins, 2020, Armah et al., 2018, Tetteh et al., 2022). Higher-income groups are more likely to have piped water connections, but even so, these sources may be unstable (Ngben and Yakubu, 2023, Zuin et al., 2011). As an alternative, residents in SSA are increasingly forced to rely on shared water sources, spending more time and money on accessing water (Chakraborty, 2022, Pierce, 2017, Sarkar, 2022). Therefore, shared water sources play an essential role in supplying water to low-income areas, as shown by studies in Kampala, Uganda (Isoke and van Dijk, 2014, Tumwebaze et al., 2023), Lilongwe, Malawi(Adams, 2018b), Lusaka, Zambia, and Cape Town, South Africa (Nel et al., 2023).

Among shared water sources, water kiosks are fixed-location facilities that may include water storage and treatment systems, where consumers can purchase water (sometimes alongside other goods). Water kiosks have a long history in Kenya. Since the colonial era, kiosks or standpipes have been used to supply water to African settlements in Kenya (Nilsson, 2011). After the 1970s, due to economic decline and concerns over self-sufficiency, these community water distribution systems were considered as an alternative formal solution for providing water in low-income areas (Nilsson, 2011, Sarkar, 2022). Boakye-Ansah et al. (2022) note that shared water sources, including water kiosks, are often managed through agreements between asset holders—such as government agencies—and intermediaries, which may include NGOs, community organizations, individuals, or private water vendors(Adams, 2018b, Contzen and Marks, 2018, Opryszko et al., 2013). In addition to the KIWASCO utility directly selling bulk water to kiosk operators within the network that it operates, Kisumu's delegated management model (DMM) incorporates water kiosks (Schwartz and Boakye-Ansah, 2023). In this model, the Kisumu Water and Sewerage Company (KIWASCO) delivers bulk water to metered Master Operators who manage DMM pipelines. These operators then sell some of this water to consumers indirectly through kiosk vendors (Nzengya, 2015).

Studies on water kiosks highlight risks of availability, affordability, quality, and accessibility. Water interruptions in kiosks occur for two main reasons: either the service provided by suppliers to the kiosks is disrupted, or kiosk operators choose to provide water only during limited hours due to cost considerations or personal commitments (Boakye-Ansah et al., 2022, Schwartz and Boakye-Ansah, 2023). A study in Malawi (Adams, 2018b) observed that, in such cases, consumers were compelled to switch to alternative water sources. Affordability is another major concern for kiosk users. Despite official pricing set by local authorities, the cost of water sold at kiosks often exceeds these regulated prices. Respondents consistently reported higher expenditures on water purchased from kiosks compared to private connections. In a field study conducted in Kisumu in 2017 and 2018, for example, water from kiosks (Ksh 0.2/litre) was priced higher than the official domestic tariff (Ksh 0.06/litre) for the first 6,000 litres of water consumed (Boakye-Ansah et al., 2022). The elevated costs are largely attributed to the presence of intermediaries, a pattern also observed in Nairobi, Kenya (Ondigo et al., 2018), Kampala, Uganda (Tumwebaze et al., 2023), and Kumasi, Ghana (Adusei et al., 2018). To address these challenges, Kisumu and other SSA cities have recently introduced prepaid dispensers (PPDs) in kiosks. Through PPDs, customers can purchase water at kiosks using tokens with pre-purchased credit. Replacing relatively unreliable operators with these automated systems is seen as a way to ensure more consistent water supply, and lower kiosk water prices (Adusei et al., 2018, Boakye-Ansah et al., 2022, Schwartz and Boakye-Ansah, 2023).

Accessibility is another key metric when assessing the efficiency and functionality of water services. The time spent collecting water is closely linked to the amount of water households can access (Devi and Bostoen, 2009). If the time required (including queuing and collection) is excessive, residents tend to reduce the frequency of water collection, and the amount of water collected is also inversely proportional to the time (Boakye-Ansah et al., 2022, Cassivi et al., 2019). Adams (2018b) and Boakye-Ansah et al. (2022) reported that queuing times in their study areas often far exceeded the time required to travel to and from the water source. Combined with the need for multiple trips per day, this imposes a significant burden on households, exacerbated by unstable supply schedules (Adams, 2018b). As women and girls primarily bear the responsibility for water collection, this issue raises concerns about gender equality and safety, especially when water must be collected at night (Rusca et al., 2017). Additionally, water quality can be impacted during collection. While studies by Zuin et al. (2011) and Tumwebaze et al. (2023) found that kiosk water quality was generally good and well-regarded, other research highlighted post-collection contamination risks. Cassivi et al. (2021), Wright et al. (2004) and Shields et al. (2015) all highlighted the deterioration in water quality between the source and stored water, implying contamination during collection, transport, or storage.

Several studies have shown that both water source types and access exhibit spatial heterogeneity (Deshpande et al., 2020, Dongzagla et al., 2022, Tetteh et al., 2022). Velzeboer et al. (2018) pointed out that the distribution of kiosks is influenced by landowners and urban chiefs and therefore does not always reflect residents' needs, often resulting in uneven coverage and further complicating water allocation practices. Lawhon et al. (2018) and Rusca and Cleaver (2022) emphasized that analysing such heterogeneity in water infrastructure can provide insights into broader socio-political dynamics. However, due to the diversity of water sources, the instability of supply, the difficulty of monitoring consumption, and issues of data quality, water usage analysis often faces significant limitations (Nauges and Whittington, 2010). Current studies on access to water points are often based either on self-reported water collection distances and times (Adams, 2018a, Adams, 2018b, Isoke and van Dijk, 2014, Kayaga et al., 2020, Tumwebaze et al., 2023, Zuin et al., 2011), which may deviate from actual distances (Crow et al., 2013, Ho et al., 2013), or on direct estimations of Euclidean distances between water points and households (Cassivi et al., 2021). Such surveys are likely to contain biases arising from sampling errors or inaccuracies in the survey instruments (Bartram et al., 2014). Additionally, interview-based studies have failed to capture the variation in water access across the urban scale.

Therefore, a city-scale quantitative accessibility analysis is necessary to examine the spatial heterogeneity of urban water distribution, an area that remains underexplored. Among studies on infrastructure accessibility, a commonly used approach is the 2SFCA method. The 2SFCA measurement originates from the gravity-based model, which captures both supply-demand dynamics and spatial distance in accessibility analysis (Luo and Wang, 2003). By introducing catchment areas to constrain the scope of accessibility, it overcomes the sensitivity of gravity-based accessibility measurements to area delineation (see Literature Review 2.5). Since its calculations are based on the spatial distribution of populations and facilities, the results not only capture spatial disparities in access but also reflect the rationality of facility distribution, making it a suitable choice for this study. The 2SFCA method and its improved versions have been widely applied in public facility accessibility research, particularly in healthcare studies (Kanuganti et al., 2016, Luo and Qi, 2009, Wang, 2012). However, to the best of available knowledge, only one study by Mahuve and Tarimo (2022) has employed the 2SFCA method in water accessibility research. Their study focused on improving the travel impedance function within the 2SFCA framework. Through a sample survey, they estimated the population within the rural wards of Dodoma Urban District in Tanzania and used this as a basis to compare their 2SFCA model with previous models. The main aim of their study was not to address the knowledge gap in the distribution of urban water resources.

Therefore, current research on water access primarily relies on field surveys, which estimate water demand based on small-scale population data. However, this approach also limits the scope of such studies, whereas large-scale infrastructure accessibility analyses typically involve the use of aggregated population data (Mizen et al., 2015, Stępniak and Jacobs-Crisioni, 2017). This aggregation method represents a unit—such as a neighbourhood—using a single point, assuming the entire population is concentrated at that point. The accessibility of that point is taken to represent the accessibility of the entire neighbourhood. This process can utilise large-scale population surfaces as the basis for aggregation, with population estimations relying on assumptions about the city. However, discrepancies between these assumptions and reality—such as variations in occupancy rates across different neighbourhoods—can lead to fluctuations in surface performance. For example, Palacios-Lopez et al. (2019) demonstrated the relationship between the covariates' quality and the population dataset's performance. They pointed out that the quality of a model that uses covariates to estimate population, such as LandScan, is affected by the availability of information. Hence, the performance of population models is constrained by the heterogeneity of study areas (Palacios-Lopez et al., 2019). This problem is particularly significant in slums, where some characteristics are often not covered by urban covariates, resulting in an underestimation of slum populations (Hanberry, 2022, Thomson et al., 2021). Thus, when population-weighted centroids are used instead of areal centroids, the population distribution affects not only demand estimates but also the location of centroids, which, in turn, impacts accessibility estimations in these areas. Considering the expansion of slums in SSA, if the 2SFCA method is to be applied to measure water accessibility in SSA urban areas, the performance of population datasets must be assessed—a factor that remains unclear.

Furthermore, in 2SFCA calculations, urban layout plays a significant role in shaping water access through the road network. The configuration of urban elements varies across regions, reflecting the factors that drive differences in infrastructure distribution within cities. Generally, variations in urban morphology are primarily identified through differences in street network topology, as street networks serve to partition urban space and facilitate material flows (Kropf, 2014, Zhang et al., 2023). The Million Neighborhoods Map (Brelsford et al., 2018) in the previous chapter uses the hierarchical structure of road networks as an indicator of neighbourhood morphological complexity, exemplifying a common approach to capturing urban-scale morphological characteristics through the use of centrality measures. These measures reveal how different areas within a city are organized and highlight material connectivity within the urban fabric (Akbarzadeh et al., 2019, Porta et al., 2012, Wang et al., 2011a, Zhao et al., 2016). Studies utilising road network centrality have identified links between urban amenities, population distribution, and infrastructure accessibility. Specifically, when network distances are used instead of straight-line distances, the topology of roads directly impacts the number of accessible facilities. However, research has yet to establish a direct link between infrastructure accessibility and urban morphology or road topology.

Therefore, to enhance the assessment of urban water services, this chapter takes water kiosks as a case study, aims to develop a workflow for applying 2SFCA to water services, and evaluates its effectiveness in quantifying water access at the city scale using secondary data. Due to data availability, kiosks in Kisumu are selected. Given the relationship between 2SFCA analysis, population data, and urban morphology, the objectives of this study are as follows:

1. Develop a workflow to evaluate geographic water accessibility using the 2SFCA method and develop recommendations for applying 2SFCA to water services.
2. Examine the spatial patterns of 2SFCA results, incorporating spatial autocorrelation analysis.
3. Assess the impact of different population surfaces on 2SFCA geographic water access measures and examine the sources of their variations.
4. Investigate the relationship between water source access and road centrality to explain accessibility in relation to urban layout.

6.2 Methodology

The study area is the urban region of Kisumu. Background information on the city's urban and water services can be found in Section 3.2. Formal water services in urban Kisumu are primarily provided by a single major service provider, KIWASCO, which directly contracts with water kiosks, except in DMM areas. To simplify the analysis, the study scope is limited to exclude small-scale community supplies and self-supply systems present in urban and peri-urban areas. The population using kiosks is estimated by excluding households with domestic piped connections based on domestic water meter density.

Measurement of infrastructure accessibility often relies on place-based approaches (see Section 2.5.1). In this chapter, an adjusted 2SFCA method was applied to measure the accessibility of water kiosks in Kisumu. The calculations were performed using different population datasets, allowing for a sensitivity analysis of how population data products affect the results. This analytical approach helped mitigate the influence of disparities between population data products on the estimation of kiosk accessibility. All three population datasets and facility distribution data are from 2020, meaning the results reflect kiosk accessibility in that year. The results were further utilized to identify patterns in kiosk access distribution and examine related factors through spatial analyses. The main workflow of the study is illustrated in the figure below (Figure 6.1).

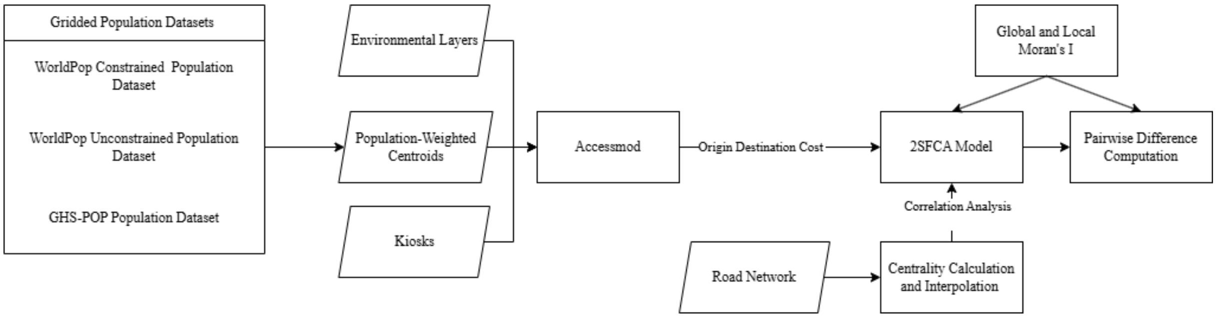


Figure 6.1 Flowchart of the major stages in kiosk accessibility analysis. Population datasets, kiosks, and environmental datasets were used as inputs for AccessMod to estimate travel time between populations and facilities, which was then applied in the 2SFCA model analysis. Following the computation of accessibility results, their spatial distribution patterns were examined.

6.2.1 2SFCA methodology and spatial accessibility analysis

The 2SFCA method was first introduced by Luo and Wang (2003) to measure spatial accessibility by calculating the supply-demand ratio for specific locations. Subsequent enhancements to the 2SFCA, such as the i2SFCA (Wang, 2018) and H2SFCA (Tao et al., 2020), adhere to similar principles (Stacherl and Sauzet, 2023) and involve two steps: calculating the facility's score (R_j) and the accessibility of population location *i* (A_i^F) (see section 2.5.2). The Gaussian function (Dai, 2010) was employed as the distance decay function as it continuously discounts accessibility, providing a more realistic representation of travel processes. When compared with various distance decay functions, the Gaussian model exhibited a higher average similarity coefficient with other models, indicating that models using this function better approximate actual accessibility (Chen and Jia, 2019). To account for variations in facility attractiveness when multiple kiosks are accessible within a travel time threshold, the Huff model (Huff, 1963) was incorporated. This model considers facility capacity and travel distance to compute the probability of a facility being chosen, assigning higher preference to kiosks located closer to the population. As distance is the primary factor in discounting accessibility, this adjustment helps to avoid unnecessary discounting in calculations, providing a more accurate representation of water collection behaviour. This adjusted 2SFCA approach, which integrates the Huff model, is also known as the 3SFCA (Liang et al., 2023). The calculation was performed in two steps：

Step1: Computation of the capacity-to-population ratio R_j for kiosk *j*

$$R_j = \frac{S_j}{\sum_{k\in[d_{kj}\leq d_0]} P_k f(d_{kj}) Prob_{kj}}$$

Step 2: Computation of accessibility A_i^F at each population location *i*

$$A_i^F = \sum_{j \in [d_{ij} \leq d_0]} Prob_{ij} f(d_{ij}) R_j$$

Where:

S_j is the capacity of kiosk *j*;

P_k is the population of centroid *k*;

d_{kj} is the distance between centroid *k* and kiosk *j* within the catchment (d_0);

$f(d_{kj})$ is the distance decay between two locations *k* and *j*, which takes a Gaussian form:

$$f(d_{ij}) = \frac{e^{-(1/2)\times(d_{ij}/d_0)^2 - e^{-(1/2)}}}{1 - e^{-(1/2)}}$$

$Prob_{kj}$ is the probability that the population-weighted centroid *k* selects kiosk *j* as the water source:

$$Prob_{ij} = \frac{S_j f(d_{ij})}{\sum_{k \in [d_{ik} \leq d_0]} S_k f(d_{ik})}$$

Similar to other accessibility measures, the results of the 2SFCA method A_i^F hold no intrinsic meaning in isolation; their significance emerges only when compared across different regions (Dalvi and Martin, 1976, Miller, 2018). Consequently, analysing the patterns of accessibility scores within the study area becomes more critical. To assess the agreement between outputs, Spearman's rank correlation coefficient and Bland-Altman plots (Bland, 1986) were employed, with the latter used to estimate and display the limits of agreement between input variables. The *blandr* package in R (Datta, 2024) was used to perform Bland-Altman analyses. Additionally, this chapter utilised the Global and Local Moran's I tools (Anselin, 1995, Getis and Ord, 1992) to analyse the spatial patterns of accessibility scores derived from the 2SFCA results. The Moran's I index is commonly used to evaluate whether a spatial feature exhibits spatial autocorrelation with its neighbouring features. In other words, it highlights the statistically significant spatial clustering patterns of accessibility values across the study area. The default distance threshold was used for the Moran's I tools to ensure that each input feature had at least one neighbour. This choice excluded other similar clustering tools, such as Incremental Spatial Autocorrelation. Global Moran's I, range from -1 to 1, representing perfect dispersion and perfect clustering, respectively. Local Moran's I, on the other hand, classifies spatial autocorrelation into hot spots, cold spots. Compared to another tool provided by ArcGIS, Optimized Hot Spot Analysis, Local Moran's I also identifies spatial outliers. After obtaining the

2SFCA outputs, both EAs with no access and those with no data were treated as zero, indicating that these EAs could not access water kiosks. The outputs were then subjected to statistical analysis. The results from different population datasets were normalised, and pairwise differences were computed. Global and Local Moran's I were used to examine both each 2SFCA output and the differences between the normalised outputs.

The relationship between accessibility and the urban environment was analysed by measuring the correlation between the topological properties of roads and the accessibility scores. Betweenness and closeness were used to measure the characteristics of road nodes in the primal graph due to their widespread application in urban studies (Jiang and Claramunt, 2004, Kirkley et al., 2018, Serra and Hillier, 2019, Shang et al., 2020, Shi et al., 2024, Wang et al., 2011a). Both metrics have been recognized as indicators of network robustness and accessibility, as well as their associations with urban layout, land use, and facility distribution (see Sections 2.4.1).

To compare accessibility and centrality within a unified framework, the centrality values obtained were smoothed using spatial interpolation techniques. The choice of interpolation method should not significantly affect the results. In this chapter, inverse distance weighting (IDW) and kernel density estimation (KDE) were used. IDW is a distance-based spatial interpolation algorithm that estimates unknown values based on weighted averages of known values within the neighbourhood. KDE, on the other hand, is a non-parametric method that estimates the probability density. Specifically, it uses the density of data within a defined range (window) to estimate the value at the window centre. In the context of centrality measurements, KDE has been regarded as effective in capturing neighbourhood characteristics (Liu et al., 2015, Liu et al., 2016). The parameters of IDW and KDE, such as power and bandwidth, produce slightly different interpolation results. However, as analysing these effects is beyond the scope of this chapter, a power of 2 was used for IDW as the default. This parameter only reflects the influence of values and is not related to any real physical process. Considering the catchment and travel scenarios of the study area (see Section 6.3.3.3), along with the search distance derived from Moran's I analysis—where each EA has at least one neighbour at approximately 1,450 meters—a bandwidth of 1,500 meters was applied for KDE. The output pixel size was set to 50 × 50 meters.

After performing the interpolation, the mean interpolated values within each enumeration area (EA) were calculated for subsequent correlation with accessibility scores. To compare outputs from different population datasets and examine the relationship between road centrality and accessibility distribution, Spearman's correlation coefficient was used to measure the relationships between accessibility maps and between interpolated road centrality and accessibility maps.

6.3 Datasets and parameterization for the study area

6.3.1 Data audit of gridded population datasets

The measurement of kiosk accessibility involves a methodological issue that requires clarification: spatial aggregation. Spatial aggregation refers to the practice of classifying individual data into geographic regions. In analysis, this means using smaller spatial units (points, lines, or areas) to represent original spatial data within the same spatial extent. In this study, the direct impact of spatial aggregation on accessibility is that different estimations of population distribution in population datasets influence the aggregation outcomes. For instance, if population data underestimate the population in a given area, the aggregation results may either retain or mitigate this bias, depending on the boundaries used for aggregation.

To evaluate the sensitivity of 2SFCA outputs to the choice of population dataset and mitigate its impact on the assessment, seven gridded population datasets available for the study area (Table 6.1) were assessed.

Table 6.1 Summary of gridded population datasets available for the study area

Dataset	Source	Spatial Resolution	Available Year (s)	Input Variables	Method	Constrained by built settlement extent
High Resolution Settlement Layer (HRSL)	Facebook Connectivity Lab and Center for International Earth Science Information Network (CIESIN) - Columbia University (2016)	30m	2015	National census data from CIESIN, binary settlement layer from DigitalGlobe imagery	Convolutional neural network for building classification, informing population distribution.	Constrained
Global Human Settlement Population (GHS-POP)	European Commission Joint Research Centre (JRC) and Center for International Earth Science Information Network (CIESIN) - Columbia University (2020)	250m/1km	1975–2030	Raw census data from CIESIN's GPW, modified using UN World Population Prospects 2019 (UN, 2020) and UN World Urbanization Prospects 2018 (UN, 2019).	Disaggregation based on built-up area distribution, classification, and density from the Global Human Settlement Layer.	Constrained
LandScan Global	Oak Ridge National Laboratory (ORNL) (2021)	1km	2000 - 2022	Sub-national census counts, spatial data, high-resolution imagery.	Multi-variable dasymetric modeling, interpolation using LandScan distribution, adjusted to geographical characteristics.	Constrained

WorldPop-Global-Constrained	WorldPop (Bondarenko et al., 2020)	100m	2020	National census data, ancillary datasets.	Semi-automated dasymetric modelling (Stevens et al., 2015), Random Forest model (RF), gridded population prediction using geospatial covariates, final adjustments aligned with UN population estimates (UN, 2020).	Constrained
WorldPop-Bespoke Country Model (WOPR) - Kenya	WorldPop (Gadiaga et al., 2023)	100m	2022	2009 and 2019 census data from Kenya's National Bureau of Statistics.	RF models combined with geospatial covariates for population estimation.	Constrained
WorldPop-Global-Unconstrained	WorldPop (2018)	100m	2000-2020	National census data, ancillary datasets.	As above WorldPop global constrained data but does not use built settlement data to constrain population	Unconstrained
Gridded Population of the World (GPW)	Center for International Earth Science Information Network (CIESIN) - Columbia University (2018)	1km	2000 - 2020	National census data from various sources (around 2010), boundary data, United Nations population estimates.	Disaggregation using a 30 arc-second grid, population distributed by land area proportion within each pixel, minimal additional geographic data (only water masks used).	Unconstrained

These gridded population datasets can be classified based on the type of modelling used to create them (Leyk et al., 2019). A dataset is 'top-down' if the model disaggregates census data or other complete population counts to cells, informed by some (or no) auxiliary data. 'Bottom-up' population datasets are those that use household survey data (i.e. small census surveys performed at local level) and ancillary data, to predict population in the cells between the surveys. All population data in the table are generated using the top-down approach. Population datasets differ in how many ancillary datasets are used to produce them. Some datasets incorporate geospatial covariates, including human settlements, built settlement extent, night light intensity, road networks, land cover and land use type, climatic factors, water features, and terrain elevation and slope (Palacios-Lopez et al., 2019) to generate weighting layers. In contrast, datasets such as GPW and GHS-POP are directly aggregated or disaggregated with very few or no covariates. Another key distinction is whether the dataset is constrained or unconstrained by built settlement extent during the aggregation or disaggregation process. Constrained models will only assign populations to settled or built-up areas, whereas unconstrained models will potentially assign a value to any cell (Thomson et al., 2022b).

6.3.2 Overview of study datasets

To ensure that the population-weighted centroids are representative, three population datasets available for Kisumu were selected to estimate the population likely to access kiosks: WorldPop Global Constrained (Bondarenko et al., 2020), WorldPop Global Unconstrained (WorldPop and CIESIN, 2018), and the GHS-POP (Schiavina et al., 2020), all of which provide population estimates for 2020. All three datasets are top-down models but differ in their degree of modelling, covariates, and methodologies (Table 6.1). Notably, both WorldPop Global Constrained and GHS-POP constrain population to areas of built settlement, meaning that they are more likely to provide accurate and detailed distributions of population. GHS-POP uses GHS-BUILT-S (GHS-BUILTS_GLOBE_R2022A, version 1.0) as its input for built-up areas, whereas WorldPop Constrained primarily relies on building footprint data from Maxar/Ecopia. Additionally, GHS-POP differs from the other two datasets in terms of its modelling approach: it either proportionally allocates population to built-up areas based on density or applies areal weighting to distribute population to non-built-up areas (Freire et al., 2016). In contrast, both WorldPop datasets use the RF algorithm (Stevens et al., 2015) to create a weighting surface for the dasymetric redistribution of census counts. Thus, WorldPop Constrained and GHS-POP differ from WorldPop Unconstrained in whether built settlements are used as a constraint, while the WorldPop datasets and the GHS-POP dataset differ in the complexity of their modelling approaches. The use of the three datasets enables an assessment of how population surfaces with varying characteristics influence accessibility analysis.

The metadata for other datasets used by the analytical methods described in Section 6.2, apart from population data, is summarised in Table 6.2 below:

Table 6.2 Characteristics of geospatial datasets used in kiosk accessibility estimation (excluding population data)

Purpose in Study	Name	Year Represented	Source	Access Type
Delineating the study area	Administrative wards in Kenya	2019	American Red Cross (https://data.humdata.org/dataset/administrative-wards-in-kenya-1450)	Open Access
	Kenya Urban Centres	2019	(Macharia et al., 2021)	Open Access
	KIWASCO service coverage	2020	Kisumu Water and Sanitation Company (KIWASCO)	By Request
Estimating population reliant on off-premises water sources	KIWASCO meter density	2020	Kisumu Water and Sanitation Company (KIWASCO)	By Request
Calculating population-weighted centroids	Kisumu Enumeration Areas	2009	Kenya National Bureau of Statistics (KNBS)	By Request
Estimating travel time	KIWASCO water facilities maps	2020	Kisumu Water and Sanitation Company (KIWASCO)	By Request

	DEM	2011	The Ministry of Economy, Trade, and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA)	Open Access
	Land cover classification	2022	FAO Water Productivity Open-access portal (WaPOR) (https://data.apps.fao.org/catalog/iso/69be3461-320f-40a6-93d7-fa4ed3db77d1)	Open Access
	Open Street Map (OSM) Road	2023/07	OpenStreetMap Foundation (OSMF) & Contributors (downloaded via https://overpass-turbo.eu/)	Open Access

The Kisumu City administrative wards map, produced and updated by the American Red Cross, contains 1,450 administrative wards across Kenya, with the version used in this study updated in April 2019. The Kenya urban centres map (Macharia et al., 2021) delineates urban centres based on a population threshold of 2,000, with boundaries digitised from 2019 data. The KIWASCO coverage map, obtained in September 2020, defines the service boundaries of KIWASCO, encompassing the available kiosks. Similarly, the meter density map, also acquired from KIWASCO, represents the spatial distribution of piped water meters, aiding in the identification of populations using household connections. The enumeration areas (EAs) map, sourced from the Kenya National Bureau of Statistics (KNBS), includes the 2009 EA boundaries, as the 2019 version was unavailable at the time. This dataset was accessed in October 2020 through the Water & Waste project, with EA-based results providing higher-resolution spatial analyses. Additionally, the KIWASCO water facilities map, obtained in September 2020, includes shapefiles of water kiosks and the pipeline network, identifying 299 kiosks (Figure 6.2). The Digital Elevation Model (DEM), extracted from the ASTER Global DEM Version 2 (NASA and METI, 2011), was chosen over Version 3 due to its enhanced void-filling and data cleanup. Land cover information was sourced from the 2022 Africa and Near East land cover classification map (FAO, 2020), based on the FAO-developed Land Cover Classification System (LCCS). The road network dataset (OpenStreetMap contributors, 2023), downloaded from OpenStreetMap, provides classifications and speed limits for roads within Kisumu City. Prior to analysis, all datasets were projected to a uniform coordinate system, Arc 1960 UTM Zone 36S, ensuring spatial consistency.

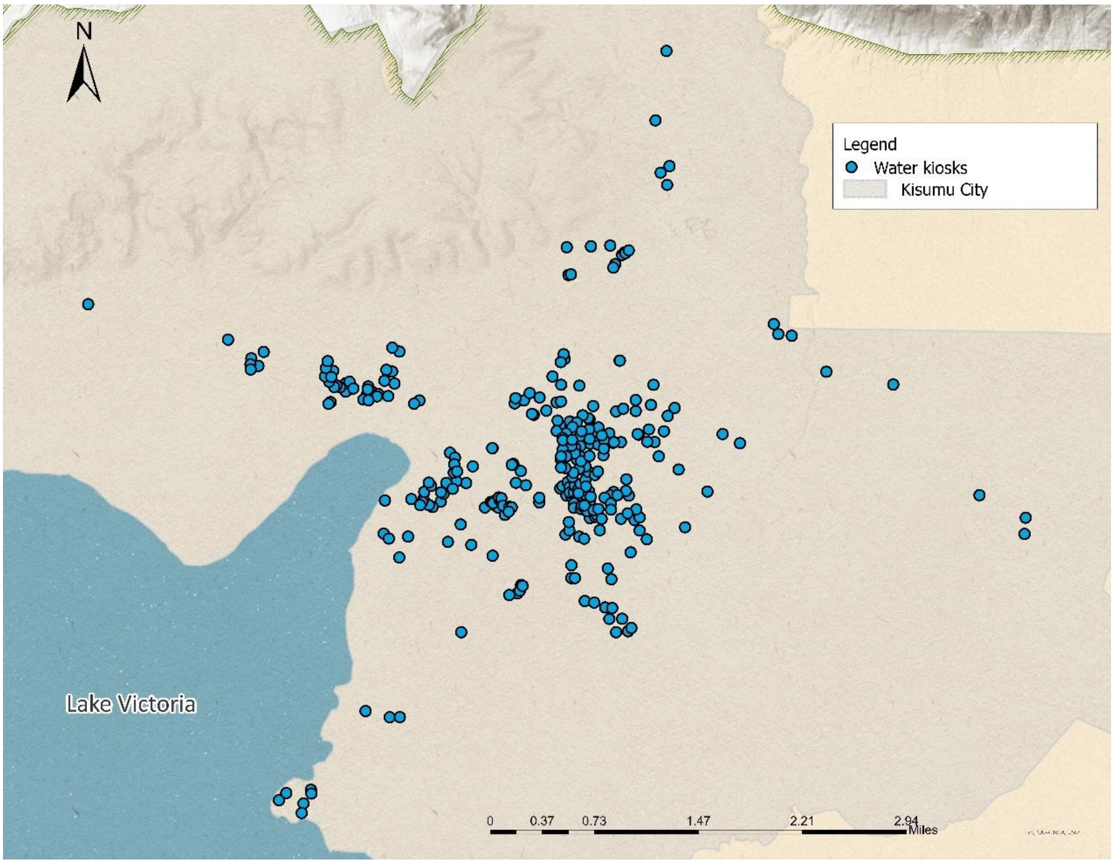


Figure 6.2 Spatial distribution of water kiosks in Kisumu.

6.3.3 Data preparation and 2SFCA parameter adjustments

6.3.3.1 Study area delineation

Narrowing the study area to the KIWASCO service coverage helps eliminate the influence of water sources other than kiosks on the accuracy of calculations. This information can be obtained from the KIWASCO coverage area map. Furthermore, since this study focuses on urban water access, the service boundary should be delineated with reference to the Kisumu City map and the Urban areas map. EAs located within the boundary were used as units for the 2SFCA analysis. The difference in map coverage also needs to be considered in this step. Since the EAs map is from 2009 and later EA boundaries were unavailable, it has a different administrative boundary from the other maps from 2019. Therefore, EAs were selected and extracted only if they were located within:

1. the 2019 boundaries of Kisumu City.

2. Kisumu's urban areas.
3. the KIWASCO coverage area.

The extracted EAs map includes 847 EAs. The average population of the units in 2009 was 389.8, and the average number of households in 2009 was 99.5, with an average area of 583,155.6 square metres of EAs. All data used subsequently was clipped to this boundary.

6.3.3.2 Spatial representation of demand for off-premises water services

In the 2SFCA analysis, the use of population-weighted centroids is considered preferable to geometric centroids in reducing spatial aggregation errors when estimating distances, particularly network distances (Mizen et al., 2015, Stępniak and Jacobs-Crisioni, 2017). Therefore, this chapter employs population-weighted centroids rather than areal centroids to calculate population locations. For data consistency, the 2020 population estimates of the WorldPop Global Constrained, WorldPop-Global-Unconstrained population, and GHS-POP datasets were used. To simplify calculations, it was assumed that the population relying on kiosks corresponds to those without household water connections. As a result, the population using household water connections was excluded from the calculation of population-weighted centroids. Using the meter density map obtained from KIWASCO, which indicates the density of households with water connections in Kisumu, and the 2019 Kenya Population and Housing Census Volume II (KNBS, 2019a), which provides the total population and number of households in each sub-location, the weighted centroids for the population accessing kiosks were derived. The calculation was carried out in the following steps:

1. Convert the sub-location household size provided by the census into a raster format.
2. Multiply the household size layer by the meter density layer using the raster calculator to generate a distribution raster for the population using household connections.
3. Subtract the household connection population raster from the overall population surface to obtain the raster for the population relying on kiosks. As the meter density distribution does not perfectly align with the population data and there may be errors in the WorldPop product, some cells in the resulting

raster had negative values. These cells were reclassified as 0, indicating that the population in these cells exclusively uses piped water.

4. Use the 'Zonal Statistics as Table' tool to calculate the population using kiosks within each EA.
5. Convert the raster data to points and assign EA names to the points based on their spatial locations.
6. Use the 'Mean center' tool to obtain the population-weighted centroids, using population values as weights.

6.3.3.3 Defining catchment areas and travel scenarios

Catchment size plays a crucial role in 2SFCA analyses. It reflects the analyst's prior knowledge of the study area and determines the model's sensitivity to spatial heterogeneity (as discussed in Section 2.5.2). The JMP framework adopts a 30-minute benchmark for round-trip access to basic improved water sources (i.e., 15 minutes one way) (WHO/UNICEF, 2018). However, in practice, water collection times in SSA vary considerably (Adams, 2018b, Cassivi et al., 2018, Dongzagla et al., 2020, Hopewell and Graham, 2014). In Kisumu, the time required to fetch water from borehole kiosks ranges from 10 minutes to over 30 minutes (Akelo and Nzengya, 2023). Therefore, earlier studies on the 2SFCA method have emphasised that catchment sizes should be adjusted according to local travel conditions (Luo and Whippo, 2012, McGrail and Humphreys, 2009), which led to the development of the variable catchment 2SFCA approach (Chen and Jia, 2019). Notably, the 15-minute water collection time (one way) refers to individuals, whereas when using population-weighted centroids instead of actual population locations, applying the same threshold would underestimate coverage and significantly reduce the number of accessible kiosks (Bryant Jr and Delamater, 2019). Consequently, this study used 30 minutes as the one-way travel time threshold.

Before estimating water collection paths, it is essential to understand the travel behaviours of Kisumu’s inhabitants. Macharia et al. (2021) reviewed transport patterns in Kenya, noting that in the capital region, 83% of trips for all purposes include walking as a mode of travel, with 41% of the trips in the city comprising walking only. In smaller cities, the reliance on walking increases significantly. Meanwhile, 65% of adults in urban slums walk to work. Although commuting travel modes differ from those used to access basic services, literature addressing this distinction is scarce. As mentioned by Watmough et al. (2022) in their study of access to healthcare facilities in Uganda, it is difficult to get information on the speeds of boda bodas or matutus (two transport tools that are also popular in Kenya). Considering that water collection is a daily activity with a low likelihood of involving transport tools, this study also considered only walking, aligning with numerous studies that use walking speed and walking distance as key parameters (Boakye-Ansah et al., 2019, Cassivi et al., 2019, Crow et al., 2013, Kim et al., 2020, Pickering and Davis, 2012).

Additionally, Watmough et al. (2022) provide insights into travel speeds across different landscapes. Their study leverages LCCS and estimates walking speeds on various land cover types and roads based on published data. These walking speed estimates were utilized to model travel scenarios.

6.3.3.4 Modelling Travel Paths with AccessMod

The distances used in access analysis should reflect real-world distances between populations and facilities (Apparicio et al., 2017, Mizen et al., 2015). Therefore, AccessMod 5 (Ray and Ebener, 2008) was used for this study to calculate network distances and travel times between locations.

AccessMod 5, developed by the World Health Organization, is widely used for analysing interactions between populations and service facilities such as healthcare centres (Hierink et al., 2023, Macharia et al., 2023). Its core functionalities include computing service areas, simulating healthcare referral pathways, and estimating facility accessibility. Unlike the Network Analyst functionality within ArcGIS, which assumes that travel occurs primarily on road networks, AccessMod simulates travel on an impedance surface, built using land cover, road, and DEM data. This surface accounts for travel speeds adjusted for terrain, slope and travel direction, enabling the calculation of travel times from any point on the map, regardless of proximity to roads. This approach is more reflective of the realities in SSA, where water collection often involves traversing varied terrain before reaching the road network. For example, residents in Malawi have complained that in addition to dangerous road conditions, they have to pass through a variety of hazardous terrain during water collection (Adams et al., 2022). A study on geographical accessibility to urban centres in Kenya also assumed a travel scenario where walking occurs across areas without road coverage (Macharia et al., 2021). In addition, since population-weighted centroids are an aggregation of the gridded population surface, they are not necessarily connected to roads. Therefore, it is reasonable that this study used Accessmod 5 in estimating travel paths.

6.4 Results

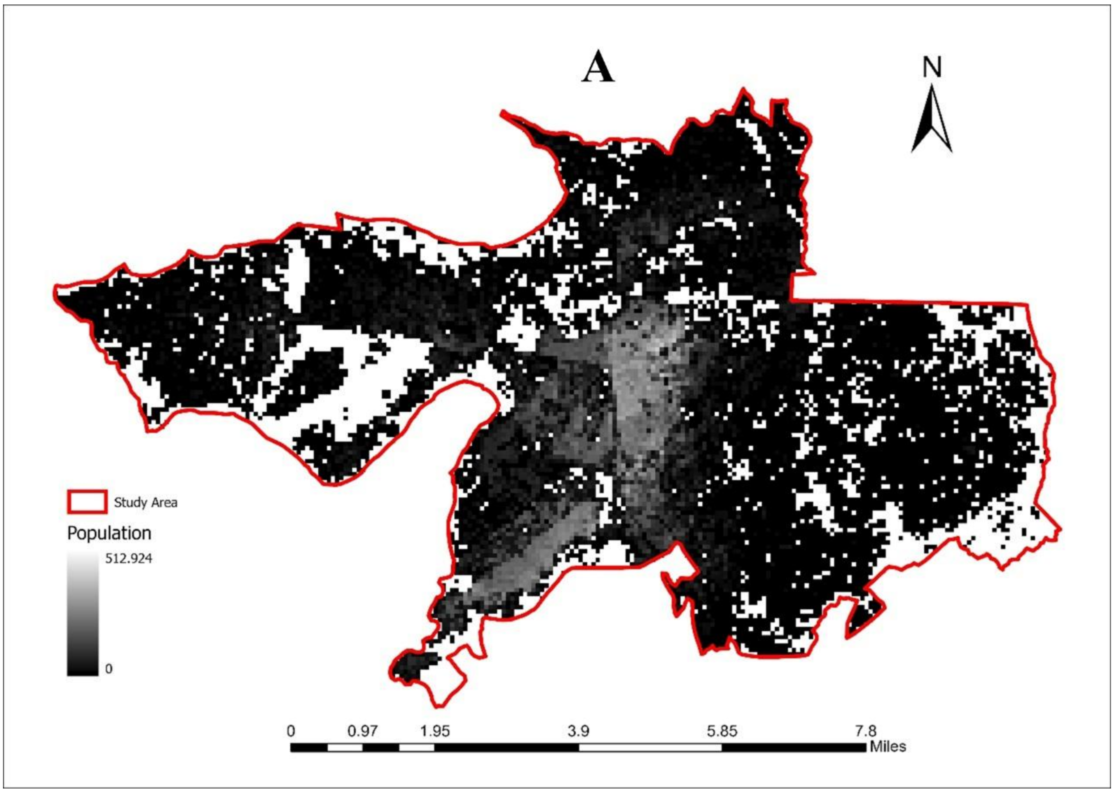
In this section, the accessibility values for the three population datasets are presented. While the population data influenced the distribution of population centroids and the calculation of demand, all other parameters and environmental factors were consistent across the three computations. This means that the differences in the results are mainly due to variations in population estimates. After obtaining the results, the commonalities and differences within the spatial outputs of the three datasets were analysed, including comparisons of their global and local Moran's I statistics, correlation coefficients, and normalized accessibility values. Additionally, the relationship between urban environments and water access was explored by examining the centrality of roads and their similarity to kiosk accessibility through Spearman correlation coefficient.

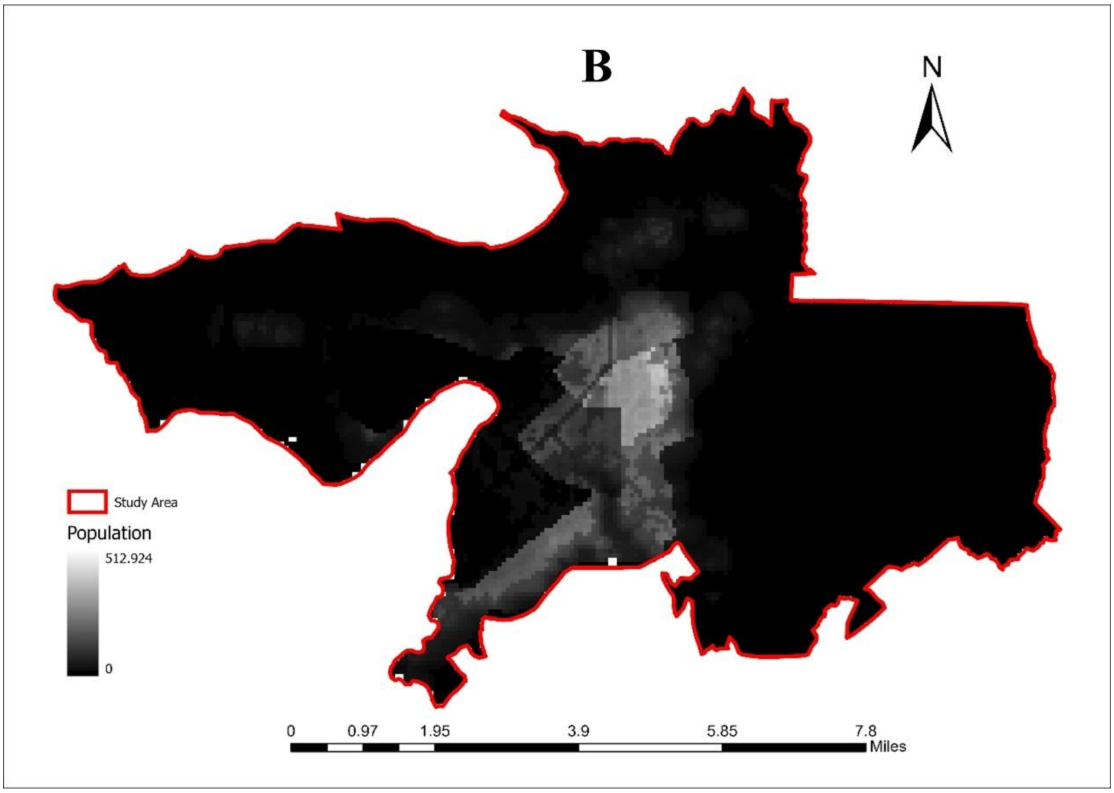
6.4.1 Differences between population products.

The GHS-POP dataset utilizes the Gridded Population of the World v4 as its data source, which uses Kenya's 2009 national census in estimating the population in the study area. This data source is consistent with those used by both the WorldPop Constrained and Unconstrained datasets. However, differences in population disaggregation methods employed by the three datasets result in variations in population density and spatial distribution estimates within the study area, as shown in Table 6.3 and Figure 6.3. The spatial distribution of high population density values is consistent across the three datasets, with concentrations observed in Kisumu’s slums (see Section 3.2.2). However, the sharpness of the boundaries defining these high-value regions varies, with the WorldPop Unconstrained dataset showing the most distinct boundaries. A greater presence of *no data* cells (hollow) can be observed in the WorldPop Constrained dataset (Figure 6.3 (A)).

Table 6.3 Summary statistics for population datasets in the study area

Population datasets	Total Population Estimates	Population Lacking On-Premises Piped Water
WorldPop Constrained	409482	340805
WorldPop Unconstrained	397346	326010
GHS-POP	457834	397123





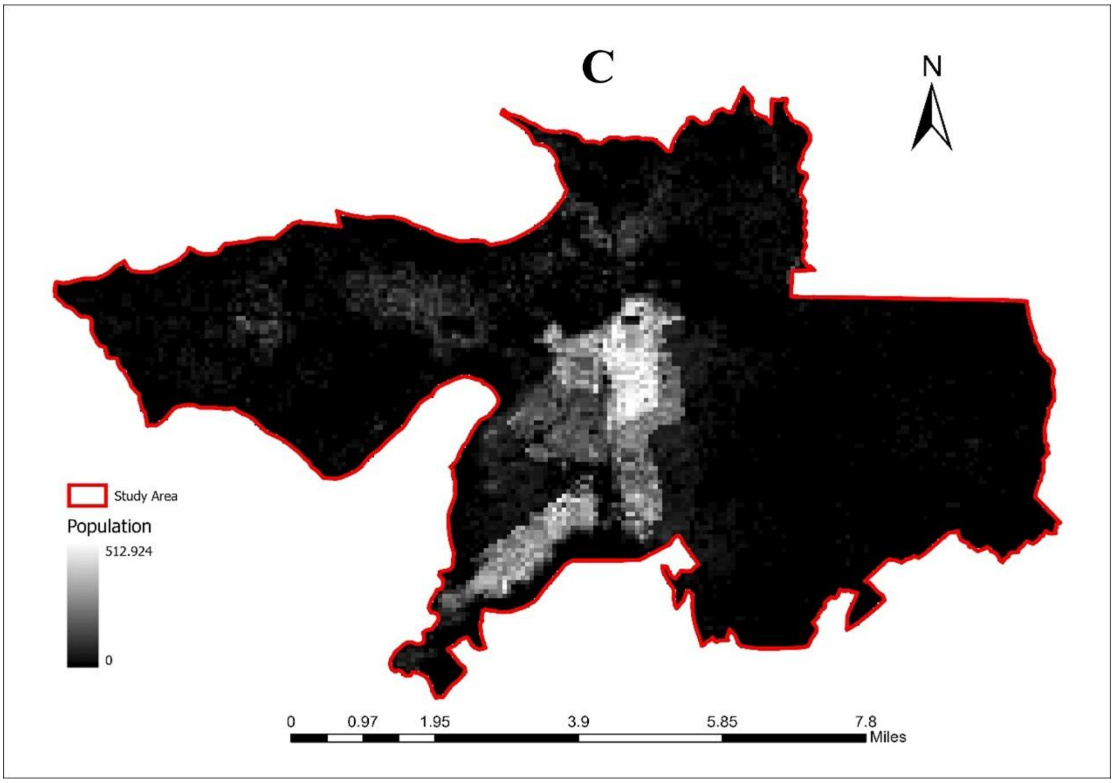


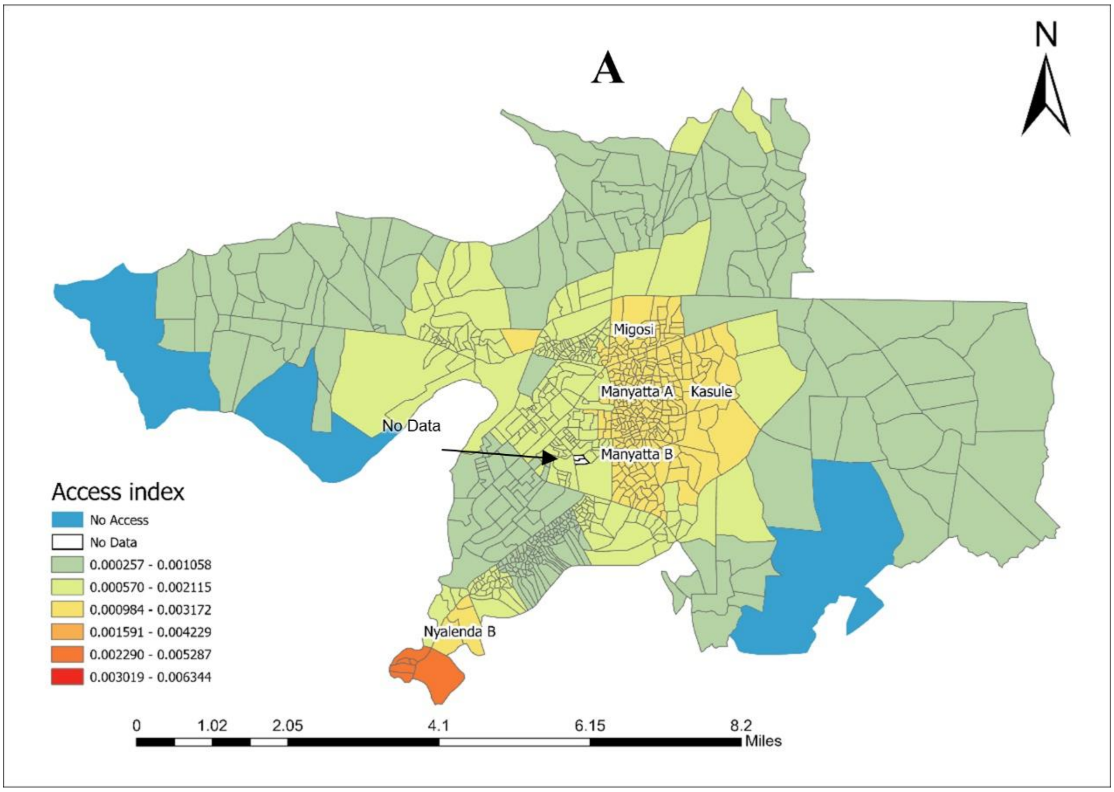
Figure 6.3 Total population estimates per grid cell in the study area from different datasets: WorldPop Constrained (A), WorldPop Unconstrained (B), and GHS-POP (C).

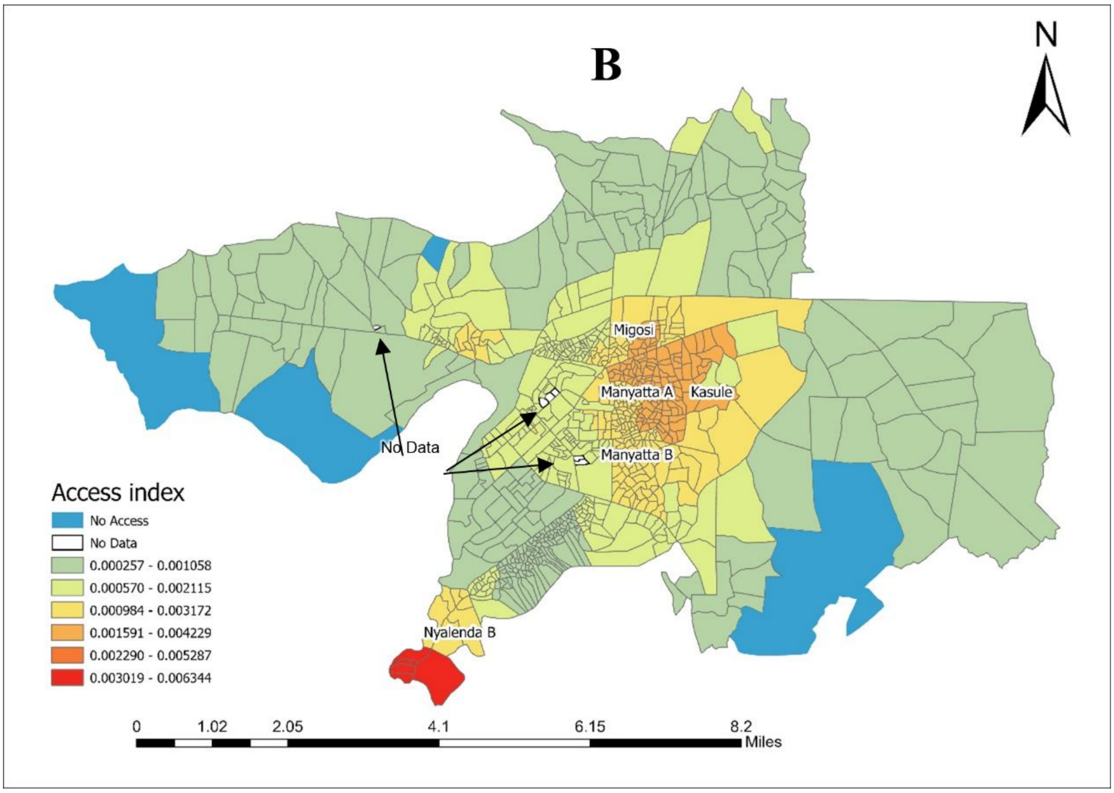
6.4.2 Mapping accessibility

The accessibility maps derived from these datasets also exhibit differences, as illustrated in Figure 6.4. "No data" indicates areas where no population is distributed, possibly due to the resolution limitations of the population datasets, while "No access" refers to areas where the population cannot access water based on the calculation. When accessibility scores for both cases are treated as zero, the unconstrained accessibility dataset shows the highest mean and standard deviation (see Table 6.4).

The figure reveals that in all three outputs, most EAs with high accessibility scores are concentrated in central Kisumu, particularly in Manyatta A, Manyatta B, Migosi, and Kasule, which host a dense network of water kiosks (Figure 6.2). It indicates heavy reliance on and easy access to these kiosks in these areas. Interestingly, the pattern aligns with the Thiessen polygon map presented in Chapter 5 (Figure 5.15), suggesting that both may be associated with the city’s spatial layout. Notably, the highest accessibility values for all three datasets are observed in Dunga EA in Nyalenda B sublocation, located in the study area's southernmost area.

Spatial differences in accessibility scores are evident among the three datasets. The WorldPop Unconstrained derived accessibility dataset tends to estimate higher values in the central region, while the WorldPop Constrained derived dataset provides high estimates for the northern and western areas. In comparison, the GHS-POP derived accessibility dataset aligns more closely with the WorldPop Constrained dataset but shows lower accessibility estimates in the central region.


A
N
Migosi
Manyatta A
Kasule
No Data
Manyatta B
Nyalenda B
Access index
No Access
No Data
0.000257 - 0.001058
0.000570 - 0.002115
0.000984 - 0.003172
0.001591 - 0.004229
0.002290 - 0.005287
0.003019 - 0.006344
0
1.02
2.05
4.1
6.15
8.2
Miles


B
N
Migosi
Manyatta A
Kasule
No Data
Manyatta B
Nyalenda B
Access index
No Access
No Data
0.000257 - 0.001058
0.000570 - 0.002115
0.000984 - 0.003172
0.001591 - 0.004229
0.002290 - 0.005287
0.003019 - 0.006344
0
1.02
2.05
4.1
6.15
8.2
Miles

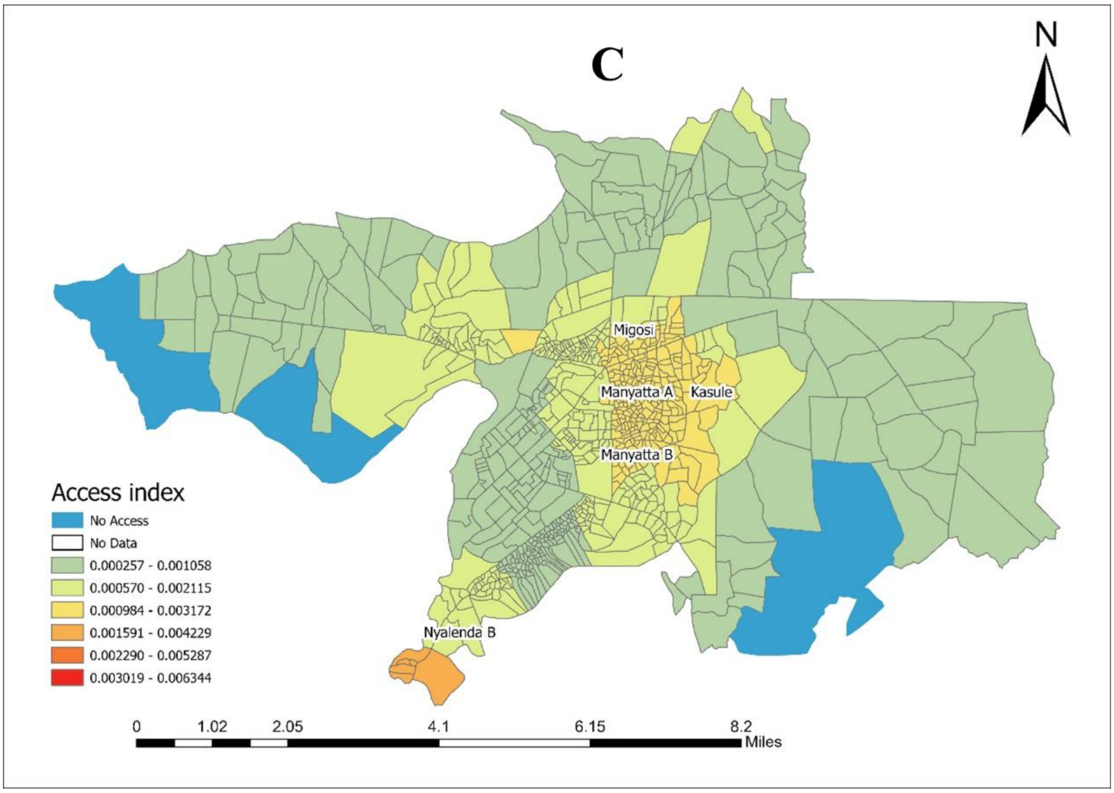


Figure 6.4 2SFCA accessibility results for EAs in Kisumu. A. WorldPop constrained dataset; B. WorldPop unconstrained dataset; C. GHS-POP dataset.

Table 6.4 Statistics of accessibility outputs for different population datasets

	WorldPop Constrained Grid	WorldPop Unconstrained Grid	GHS-POP Grid
Mean	0.0016	0.0018	0.0015
Median	0.0017	0.0017	0.0015
Std Dev	0.0010	0.0012	0.0008
Maximum	0.0046	0.0063	0.0038

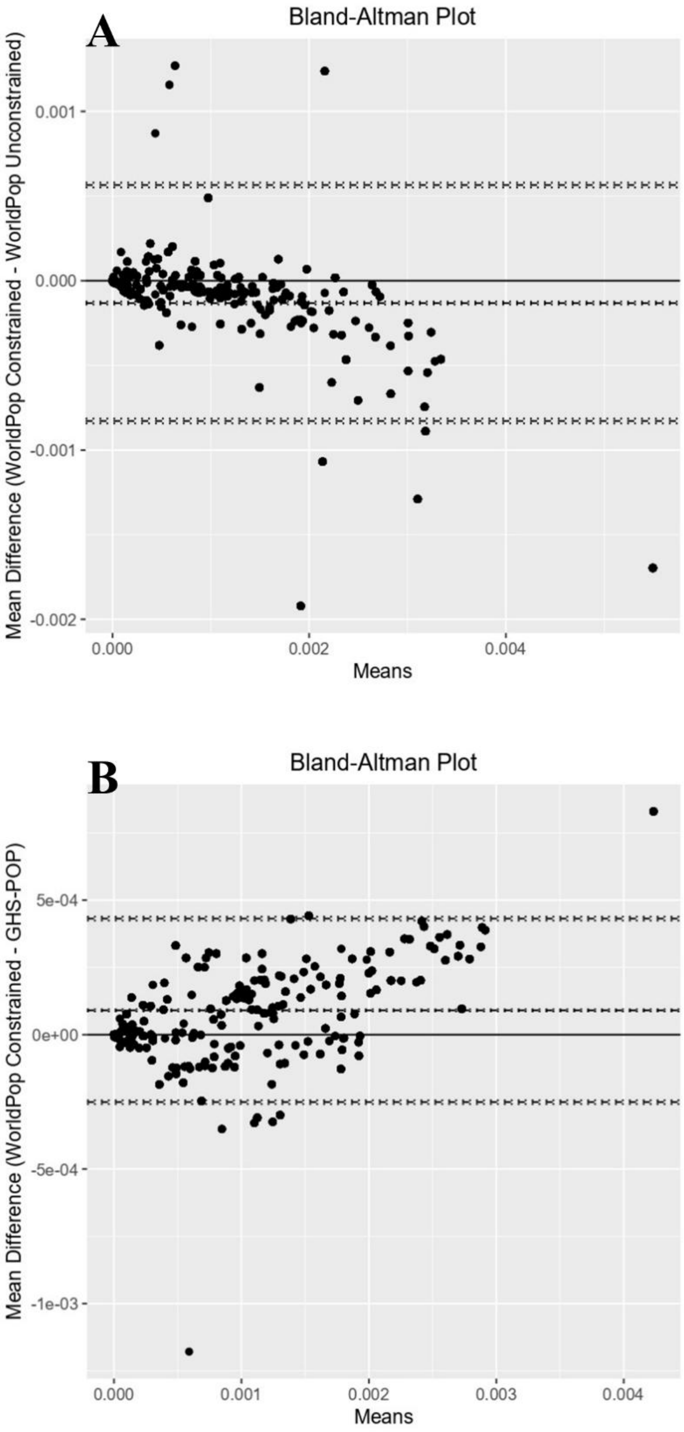
6.4.3 Comparison of accessibility across population datasets

The study also conducted pairwise comparisons of the three outputs, calculating the correlation coefficients between them, as well as the Global and Local Moran's I indices of the differences between the normalized accessibility scores. While the former illustrates the similarity between the outputs, the latter reveals the distribution of their differences. Table 6.5 shows that the three population datasets generally exhibit high similarity. However, the unconstrained output demonstrates lower similarity to both the WorldPop constrained and GHS-POP outputs, which is consistent with the previously observed results.

Table 6.5 Spearman correlations between accessibility outputs (derived from 847 EAs, p values <0.001).

Output		Spearman
WorldPop Constrained Grid	WorldPop Unconstrained Grid	0.9809
WorldPop Constrained Grid	GHS-POP Grid	0.9862
WorldPop Unconstrained Grid	GHS-POP Grid	0.9679

The Bland-Altman plot provides a visual assessment of the agreement between accessibility outputs (Figure 6.5). The X-axis represents the EA-level mean of each pair of input data, while the Y-axis shows their difference. The plots below are based on a 95% confidence interval. The mean differences across the three comparisons are relatively small (close to 0); however, notable variations exist in the upper and lower limits of agreement. Specifically, the WorldPop Constrained and WorldPop Unconstrained pairs, as well as the WorldPop Unconstrained and GHS-POP pairs, exhibit similar ranges for their upper and lower limits, whereas the WorldPop Constrained and GHS-POP pair has a much narrower range. This suggests a higher level of agreement between the WorldPop Constrained and GHS-POP datasets. Additionally, the distribution of points indicates that the differences between the WorldPop Constrained and GHS-POP outputs are more evenly spread compared to the other two pairs.


A
Bland-Altman Plot
Mean Difference (WorldPop Constrained - WorldPop Unconstrained)
0.001
0.000
-0.001
-0.002
0.000
0.002
0.004
Means
B
Bland-Altman Plot
Mean Difference (WorldPop Constrained - GHS-POP)
5e-04
0e+00
-5e-04
-1e-03
0.000
0.001
0.002
0.003
0.004
Means

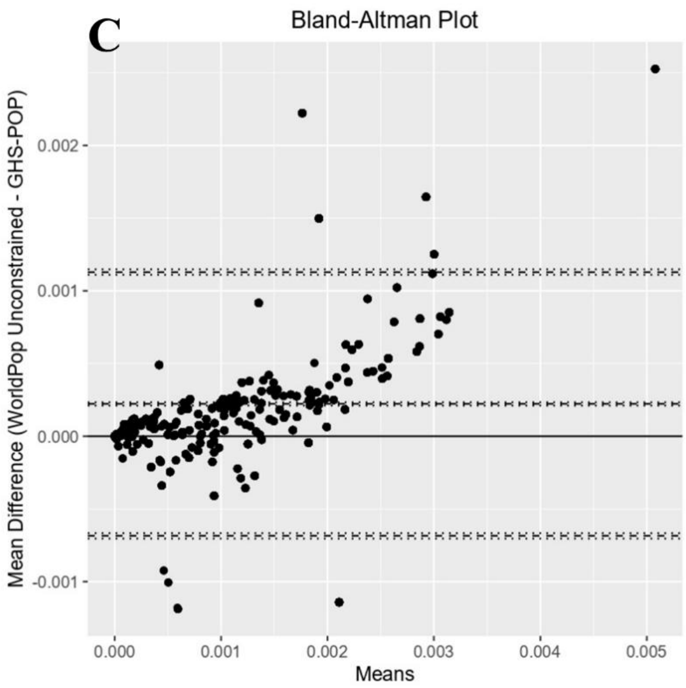
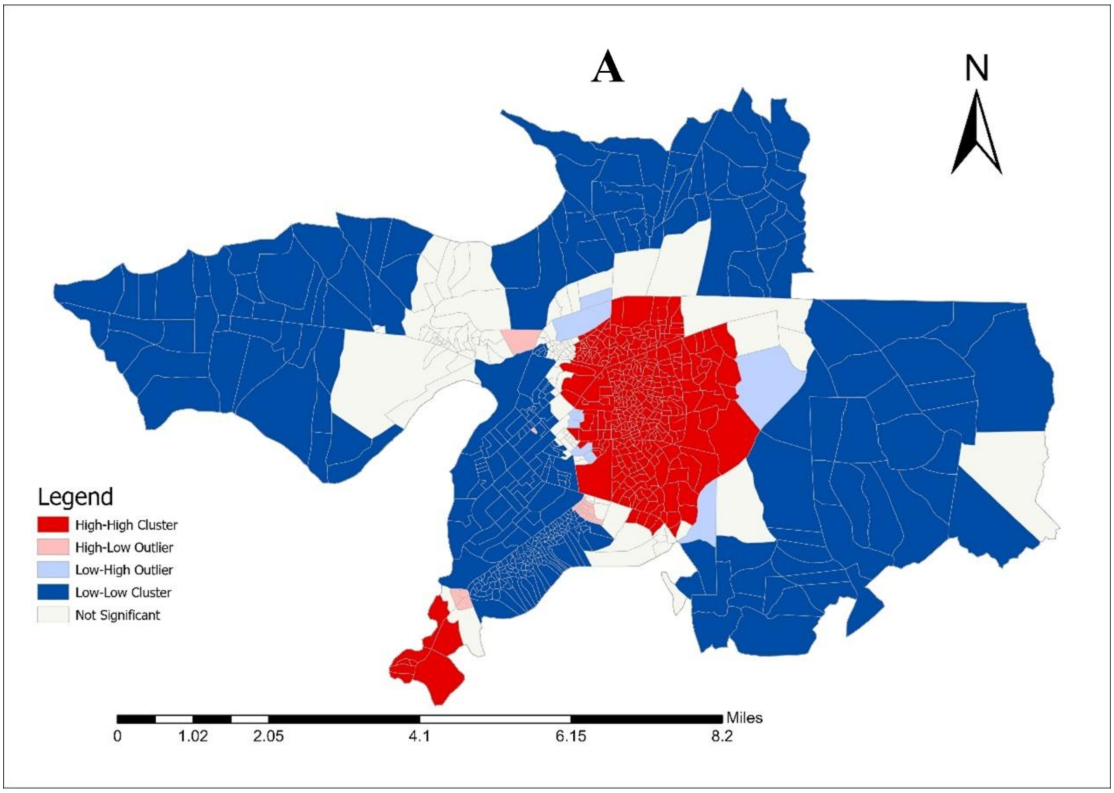
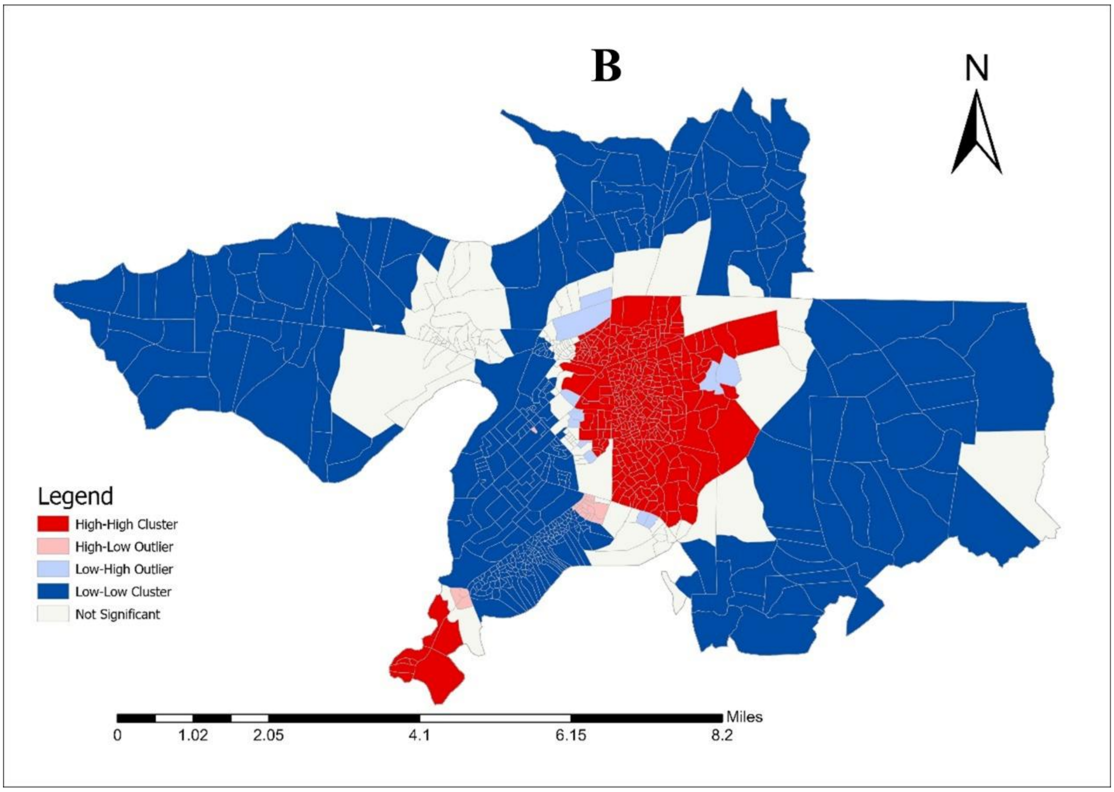


Figure 6.5 Bland-Altman plots comparing accessibility scores by Enumeration Area. (A) WorldPop Constrained vs. WorldPop Unconstrained, (B) WorldPop constrained vs. GHS-POP, and (C) WorldPop Unconstrained vs. GHS-POP. The significance level is set at 0.95. The three dashed lines represent, from top to bottom, the upper limit of agreement, the average difference, and the lower limit of agreement.

Spatial autocorrelation analysis further revealed the geographic distribution of kiosk accessibility (Figure 6.6). Global Moran's I results indicate significant clustering of kiosk accessibility within Kisumu, although the degree of clustering varies slightly across the three 2SFCA outputs. The order of clustering intensity, from highest to lowest, is WorldPop constrained (0.9077), GHS-POP (0.9039), and WorldPop unconstrained (0.8855). All results have p-values far below 0.001 and z-scores exceeding 96.

The Local Moran's I analysis, on the other hand, corroborates the locations of high accessibility values observed in the previous section. It clearly illustrates a pattern where kiosk accessibility decreases from the centre of Kisumu outward.


A
N
Legend
High-High Cluster
High-Low Outlier
Low-High Outlier
Low-Low Cluster
Not Significant
Miles
0
1.02
2.05
4.1
6.15
8.2


B
N
Legend
High-High Cluster
High-Low Outlier
Low-High Outlier
Low-Low Cluster
Not Significant
Miles
0
1.02
2.05
4.1
6.15
8.2

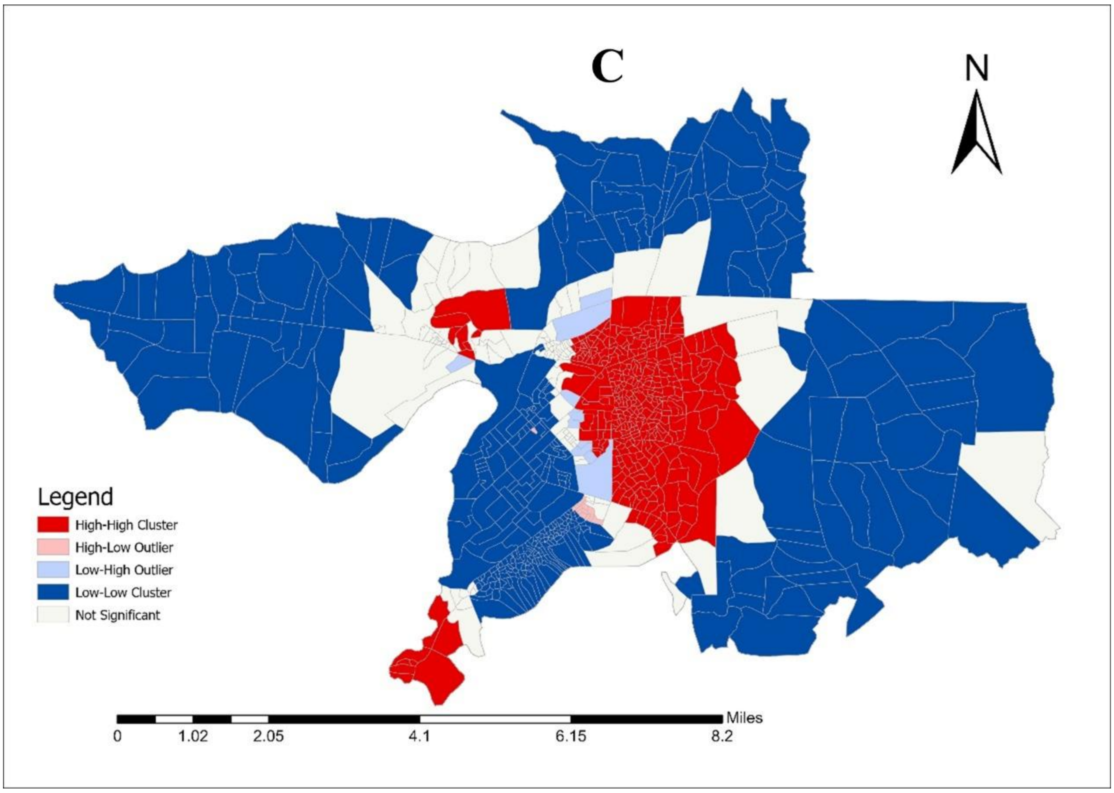
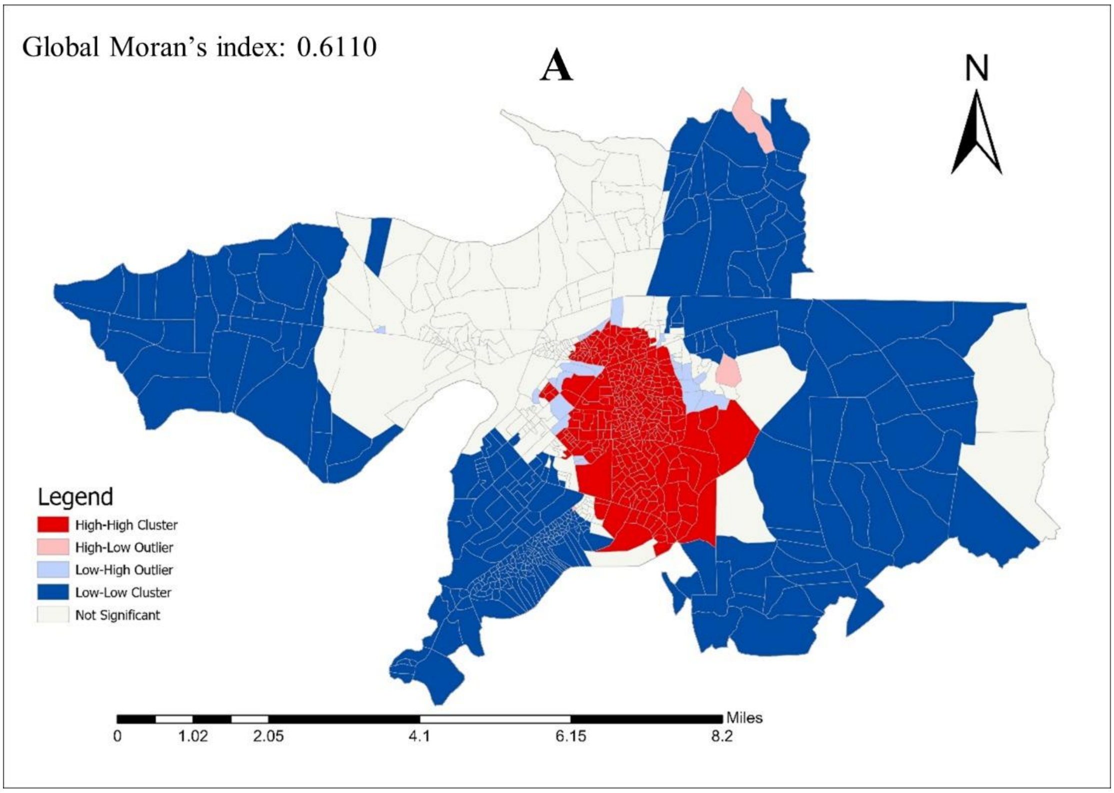
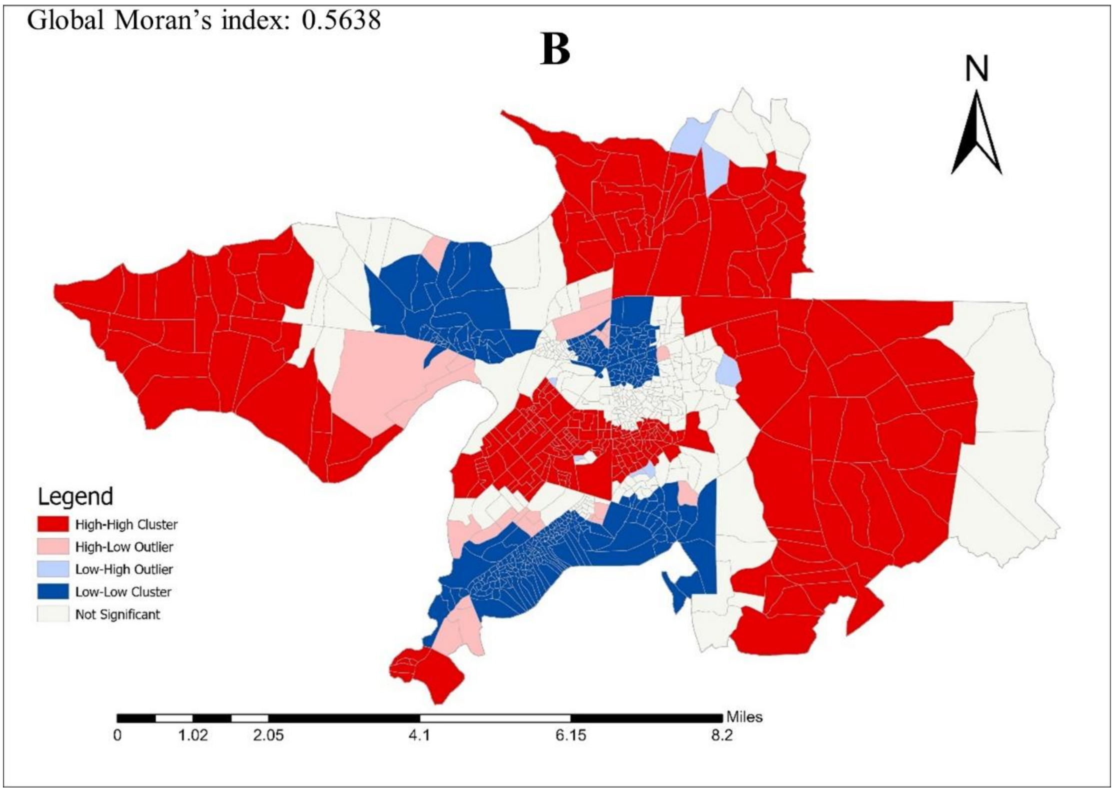


Figure 6.6 Global and local Moran's I of accessibility for EAs in Kisumu. A. WorldPop constrained dataset; B. WorldPop unconstrained dataset; C. GHS-POP dataset. It shows whether areas with high or low accessibility scores are surrounded by similar regions.

The Global Moran's I index of differences reveals that the differences between the WorldPop Unconstrained accessibility dataset and both the WorldPop Constrained and GHS-POP datasets are pronounced with stronger clustering. This suggests that the population estimates in the unconstrained population dataset likely differ significantly in their spatial distribution from the other two datasets. The study further normalised the 2SFCA results and calculated the pairwise differences, followed by Moran's I analysis of these differences, as shown in Figure 6.7. The differences between the constrained and unconstrained accessibility datasets are primarily observed in central urban areas, where accessibility is higher. In contrast, both datasets exhibit considerable agreement in the more remote areas of Kisumu. On the other hand, the differences between GHS-POP and the other two datasets are concentrated in areas with lower accessibility. However, GHS-POP and unconstrained accessibility datasets demonstrate significant agreement in the central areas.


Global Moran's index: 0.6110
A
N
Legend
High-High Cluster
High-Low Outlier
Low-High Outlier
Low-Low Cluster
Not Significant
Miles
0
1.02
2.05
4.1
6.15
8.2


Global Moran's index: 0.5638
B
N
Legend
High-High Cluster
High-Low Outlier
Low-High Outlier
Low-Low Cluster
Not Significant
Miles
0
1.02
2.05
4.1
6.15
8.2

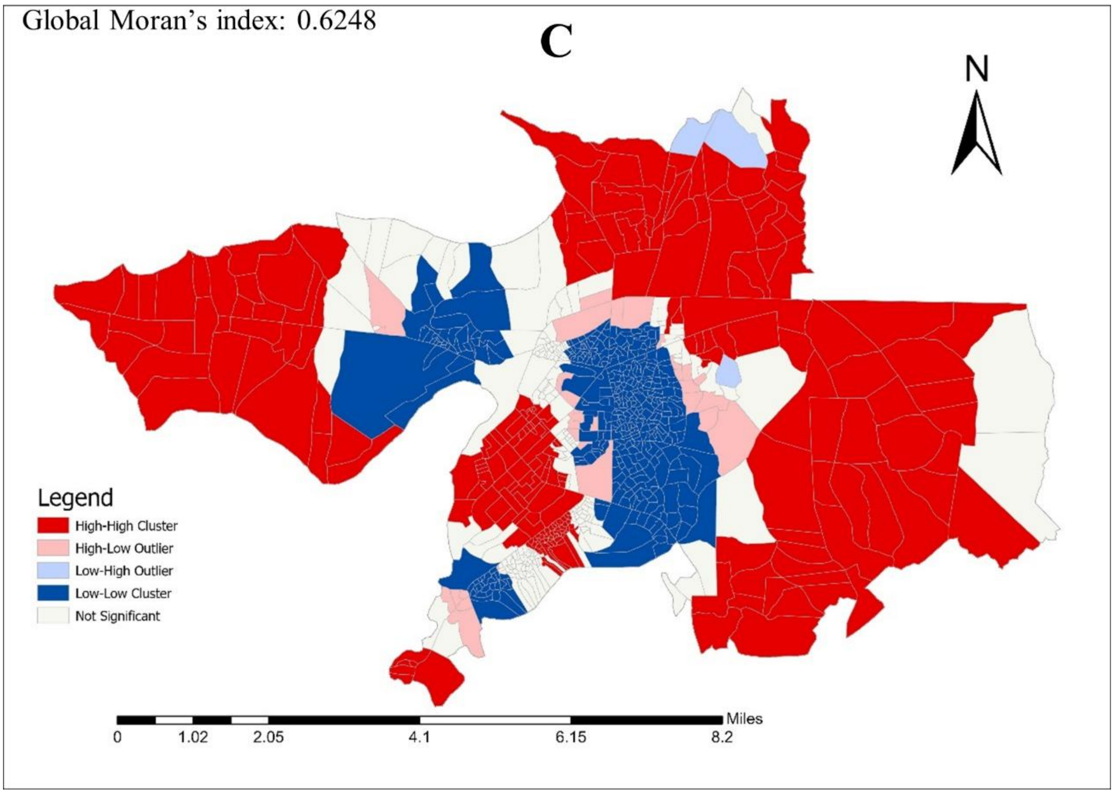


Figure 6.7 Global and local Moran's I of the differences in normalized accessibility scores between population datasets. A. WorldPop Constrained vs. WorldPop Unconstrained; B. WorldPop Constrained vs. GHS-POP; C. WorldPop Unconstrained vs. GHS-POP.

6.4.4 Correlation analysis of road centrality and kiosk accessibility

The analysis of road centrality within Kisumu reveals that areas with high road centrality overlap to some extent with the distribution of kiosk accessibility, while also extending into the central urban area (industrial area). An exception is observed in the IDW results of betweenness centrality, where high values roughly form the shape of the road network. This pattern is likely due to the nature of betweenness, which reflects the number of shortest paths passing through a node and is closely related to traffic efficiency. Consequently, betweenness values are concentrated along major roads and decay rapidly in the IDW interpolation. This rapid decay limits the influence of high betweenness on surrounding areas, leading to a lower correlation coefficient with kiosk accessibility in the analysis (Table 6.6).

Additionally, while both closeness and betweenness KDE interpolations overlap the same two regions with dense kiosk distributions (Figure 6.8), their distributions exhibit distinct differences. Betweenness KDE features two core areas of concentration, thus demonstrating a stronger correlation with 2SFCA accessibility compared to closeness KDE interpolation, as shown in Table 6.6. Moreover, regardless of the interpolation method, the correlation between road centrality and the population datasets consistently follows the order: WorldPop Constrained > WorldPop Unconstrained > GHS-POP.

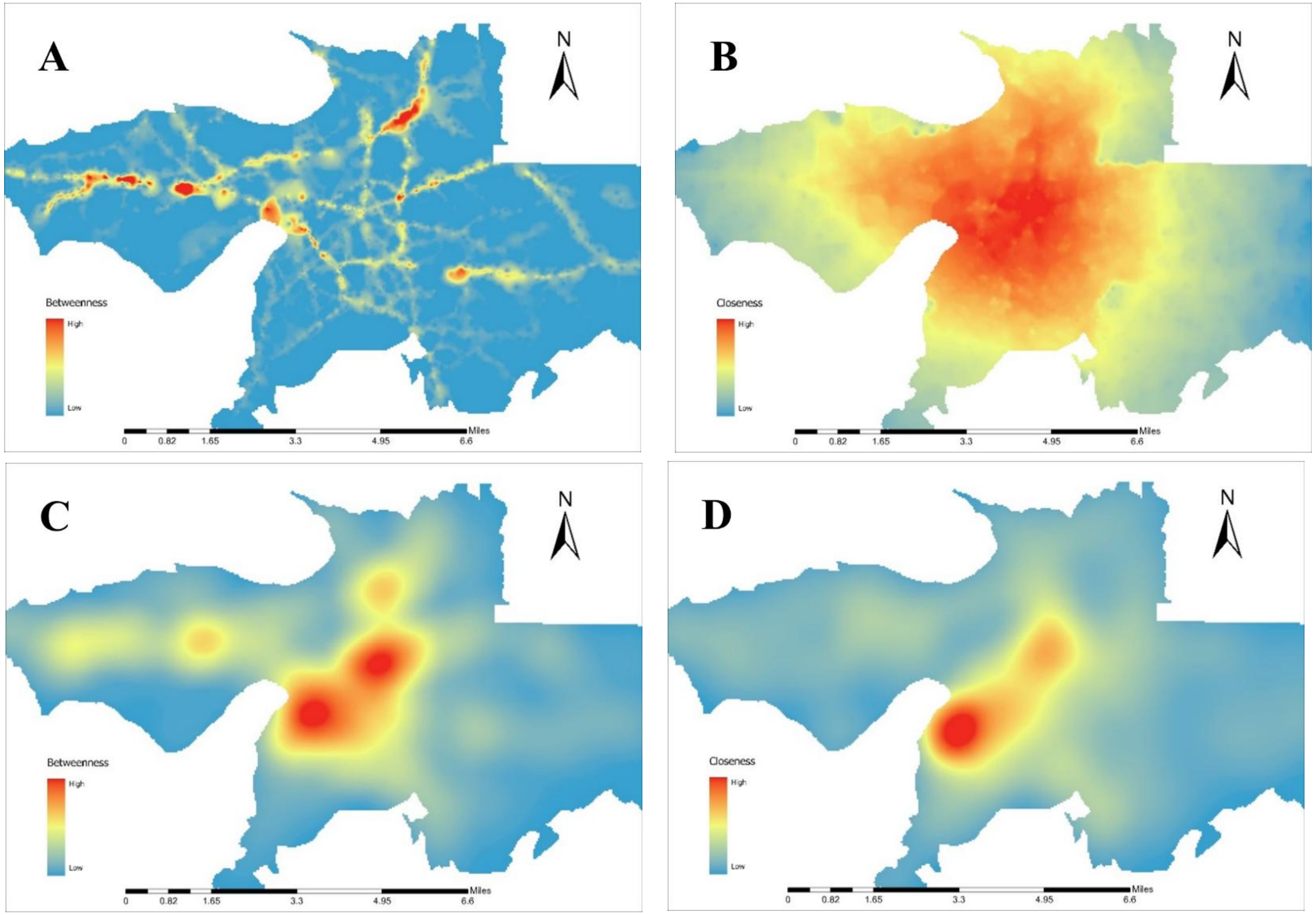


Figure 6.8 Interpolation of road betweenness and closeness using IDW and KDE. A. IDW Interpolation of betweenness; B. IDW Interpolation of closeness; C. KDE Interpolation of betweenness; D. KDE Interpolation of closeness.

Table 6.6 Spearman correlation coefficient (derived from 847 EAs, p values <0.001).

	Kernel		IDW	
	Betweenness	Closeness	Betweenness	Closeness
WorldPop Constrained	0.6147	0.5390	0.3473	0.6386
WorldPop Unconstrained	0.6002	0.5370	0.3258	0.6150
GHS-POP	0.5637	0.4931	0.3139	0.5962

6.5 Discussion

Due to the widespread challenges of urbanisation faced by SSA countries, shared water sources play a critical role in urban water supply systems. Investigating the accessibility of water kiosks provides valuable insights into the heterogeneous distribution of water resources in southern city environments and contributes to monitoring progress toward SDG 6. This study employed the 2SFCA method to analyse the spatial distribution of accessibility to water kiosks within the study area. Widely applied in healthcare research, this method has been shown to effectively capture the distribution of facilities and populations, as well as the social factors associated with accessibility (Chen and Jia, 2019, Kanuganti et al., 2016, Luo and Qi, 2009, Wan et al., 2012, Wang, 2012). However, it has not yet been applied to urban water accessibility assessments. Building on previous 2SFCA studies in healthcare and secondary data on the study area and SSA cities, the 2SFCA model was configured to account for population aggregation, catchment and travel scenarios. The results derived from WorldPop and GHS-POP datasets consistently show that EAs with high values are concentrated in regions where kiosks are most densely distributed, with accessibility declining outward from this central zone (Figure 6.4).

Further analysis differentiated the outputs produced using different population datasets. The significant differences emerge between the outputs generated from WorldPop Unconstrained data and those derived from the constrained datasets (WorldPop Constrained and GHS-POP data). WorldPop Constrained and GHS-POP exhibit higher spatial similarity in their 2SFCA estimates, with more comparable clustering patterns (Figure 6.6). However, accessibility scores derived from GHS-POP population data are generally lower. Analysis of the disparities reveals that the differences in accessibility scores between constrained and unconstrained datasets appear in similar regions (Figure 6.7).

In centrality analyses, nodes with high betweenness typically align with roads characterized by high transport efficiency, whereas nodes with high closeness represent convenient transportation hubs (Casali and Heinimann, 2019, Lan et al., 2022). These metrics are closely linked to urban planning and land-use types. In the study area of Kisumu, correlation analyses revealed a positive relationship between road betweenness and closeness centrality and the 2SFCA results (Table 6.6). The IDW interpolation of closeness has a spatial distribution similar to 2SFCA accessibility, with a decay pattern radiating outward from regions with high values. Furthermore, KDE interpolations of both betweenness and closeness cover two areas densely populated with kiosks. Rui and Ban (2014) noted that built-up areas and urban greenbelts are positively associated with various centralities. Similarly, Mengistie et al. (2023) highlighted that urban socioeconomic attributes, such as walkability, block size, and GDP, positively correlate with both betweenness and closeness. Overall, road centrality tends to decrease with increasing distance from the central business district (CBD). Shi et al. (2024) emphasized the strong relationship between closeness centrality and the distribution of residential life facilities, noting that road centrality exerts a greater influence on the distribution of urban facilities than population distribution. Liu et al. (2015) further pointed out that, beyond betweenness, the density of road centrality positively correlates with road density. As centrality density increases, construction land expands and becomes more compact. From these, three main inferences can be drawn:

1. Areas with higher 2SFCA values are concentrated in the central study area, including Manyatta A, Manyatta B, Migosi, and Kasule. These areas are characterized by a dense distribution of kiosks and high centrality values. Therefore, the observed high accessibility in these areas is likely not coincidental, as kiosks are primarily intended to serve residents who lack access to household piped water. These high 2SFCA areas are located within Kisumu's slums (Okotto-Okotto et al., 2015), where the high population and building densities are more likely to lead to high road centrality values and densities. However, it should be noted that the correlation between the KDE of closeness and accessibility is lower than that of the KDE of betweenness. Visually, high values of closeness are concentrated in the industrial area located in Bandari. Due to the dense concentration of low-income households residents in this area, it has been classified as a slum by UN-Habitat (2005), (2020). This suggests that in Manyatta and its surrounding areas, where kiosk accessibility is high, the contribution of road patterns to accessibility is limited. Instead, the high accessibility in these areas is primarily attributed to the density of major urban roads with high betweenness.
2. Population datasets predict population distributions differently, leading to variations in results. The density and spatial distribution differences among population datasets stem from variations in their disaggregation methods. In addition to the significant differences in population distribution estimates between the constrained and unconstrained datasets due to the use of built-up areas, the limitations of the datasets

also include the accuracy of auxiliary data for built-up areas in GHS-POP (Yin et al., 2021), and the resolution of input variables (e.g., nighttime lights and land cover types) in WorldPop's random forest approach (Thomson et al., 2022a). In a further analysis based on slum populations, Breuer et al. (2024) also pointed out that the quality of input census data may affect the accuracy of population estimates. In accessibility calculations, the estimation of demand and travel distance depends on the spatial and numerical accuracy of population datasets, which affects the results. For example, a study on gridded population data's impact on healthcare accessibility in SSA highlighted that population data characteristics directly influence the estimated number of people covered by healthcare facilities (Hierink et al., 2022). Since road betweenness and closeness centrality are related to travel efficiency, and the distribution of kiosks directly influences the travel distance for water access, both are key factors that affect the distribution of accessibility. Based on the maps of the two factors and the accessibility outputs, as well as the correlation coefficient between road centrality and accessibility outputs, the WorldPop Constrained dataset likely more accurately reflects the population disparities within the study area. This finding aligns with other studies on population dataset characteristics (Kuffer et al., 2022, Thomson et al., 2021).

3. In the study area, kiosks partially meet the water needs of residents in informal settlements. This is primarily due to the dense distribution of kiosks within informal settlements and the construction of highly connected urban roads. However, residents in certain slums still face difficulties in accessing water. On the other hand, previous research also indicated that queueing times in certain slums, such as Manyatta and Migosi, are longer than in others, such as Nyalenda (Sima et al., 2013). Based on this observation of kiosk capacity and the findings on kiosk accessibility, it is recommended that the service capacity of kiosks be increased in high-accessibility areas such as Manyatta and Migosi. Additionally, improving connectivity between informal settlements in the Nyalenda area and other regions should be prioritized. In other words, in terms of water access, improvements in overall layout should be considered after the development of arterial roads.

The 2SFCA workflow established in this study effectively captured the spatial characteristics of urban water access. The travel mode for water collection in the study area was set as walking, as noted by Crow et al. (2013) and Kim et al. (2020), who observed that residents in Kenya typically fetch water on foot. Based on this scenario, a 30-minute catchment (one-way trip distance threshold) was applied. Previous research has highlighted that catchment size in 2SFCA analysis should vary by region, as excessively large or small catchments fail to capture accessibility patterns (Chen and Jia, 2019, Luo and Whippo, 2012). A secondary reason for adopting a 30-minute threshold instead of the 15-minute standard recommended by the JMP is the findings of Akelo and Nzengya (2023) , which indicate significant variation in water collection times from kiosks in Kisumu. A more fundamental reason is that the JMP's 15-minute threshold refers to household-level water access, whereas in this study, EAs rather than household-level population inputs were used. Applying the original threshold could exclude some kiosk destinations that would otherwise be accessible. To account for this, the catchment area was expanded while introducing a Gaussian distance decay function to appropriately discount accessibility scores. This adjustment provides a more realistic representation of local water-fetching behaviours. The inclusion of the Huff model further enhances this representation. As a place-based accessibility model, 2SFCA emphasises the role of distance in accessibility; therefore, weighting accessibility scores based on kiosks' distance to consumers is a more reasonable and practical approach. The results confirm that these localised adjustments are appropriate. The workflow can be transferred to areas where water points are well mapped or applied to broader water accessibility measurements by incorporating predictive techniques for water points (Yu et al., 2019b).

Beyond indicating the relationship between accessibility and urban morphology, the similarity between road centrality and accessibility also suggests that, at least from the perspective of travel efficiency, the WorldPop constrained dataset performs better within the study area. However, this conclusion requires careful consideration. The accuracy of population estimates in urban areas with slums may fluctuate due to the unique characteristics of these settlements. Comparative studies in South Asia have shown that GPW, GHS-POP, WorldPop, and LandScan exhibit varying errors across countries, particularly in areas with high population densities and rapid population growth (Yin et al., 2021). Consequently, all population datasets tend to underestimate the population in slums (Breuer et al., 2024, Thomson et al., 2021, Thomson et al., 2022a). Given that kiosks' users are expected to be densely concentrated in slums, all selected population datasets likely overestimate kiosk accessibility. This underestimation of population arises primarily from the inability to accurately delineate slum boundaries and update population distribution patterns accordingly (Breuer et al., 2024, Kuffer et al., 2016, Thomson, 2020). With more precise slum delineation, this issue can be mitigated, thereby improving the accuracy of the WorldPop constrained dataset (Thomson et al., 2022b). Therefore, in well-established slums, where buildings are adequately identified, this should not pose a significant problem. Population maps (see Figure 6.3) indicate that slum populations in Kisumu are effectively captured, which likely explains the superior performance of the constrained dataset in this context. However, this finding may not be generalizable to other cities, as other SSA cities may have newly emerging slums, where the advantages of the constrained dataset are less pronounced. Moreover, given the factors influencing the performance of population datasets as discussed earlier, it is essential to examine the data sources and covariates when applying these population products in other urban contexts.

The current study and its methodological configuration have some limitations. The observed decline in accessibility from the centre outward is partly attributable to the disadvantaged positional status of population centroids located at the periphery of the study area when calculating accessibility. Therefore, catchment size should be adjusted to account for this effect. Moreover, the population in the study area was divided into two groups: those using household piped water and those using kiosks. However, this method oversimplifies the water supply situation in Kisumu. On one hand, it is common in developing countries for households to use multiple water sources (Daly et al., 2021, Nauges and Whittington, 2010), meaning that those with access to household piped water may also use water kiosks. On the other hand, urban residents in SSA often supplement their water supply with alternative sources. The study did not account for other shared water sources in Kisumu, such as self-supply water sources. Consequently, the estimation of demand for kiosks in the 2SFCA analysis does not fully capture the complexities of water usage patterns, nor does it entirely reflect the accessibility of shared water sources. Additionally, the analysis is based on the distance and travel time between facilities and populations. While this approach provides a broader picture of urban water fetching compared to small-scale, field-based studies, it overlooks factors such as time spent on water purchase, pumping, and queuing. Although this does not impact urban-scale infrastructure planning, it implies that the 2SFCA results may not effectively reflect residents' actual water-fetching experiences.

It is worth noting that this study is also affected by the Modifiable Areal Unit Problem (MAUP) resulting from spatial aggregation. Spatial aggregation is often employed for three main reasons: (1) certain patterns can only be revealed at specific scales (Heuvelink, 1998, Marceau et al., 1994, Seyfried and Wilcox, 1995); (2) data for the study area may be incomplete or anonymized to protect privacy; and (3) spatial aggregation reduces computational demands. In accessibility analyses, aggregation is frequently applied to population data when measuring distances between populations and amenities (Chen, 2019, Fransen et al., 2015, Kiani et al., 2021, Xing et al., 2020), primarily for the third of these reasons. Origin-destination matrices grow exponentially with the number of units, increasing calculation times. In the 2SFCA analysis presented in this chapter, EAs, which are considered relatively detailed, were used as aggregation units for population-weighted centroids. This approach cannot fully overcome the errors introduced by the MAUP (Bryant Jr and Delamater, 2019, Wong, 2009). Compared to individual-level data, the use of EAs tends to overestimate accessibility, particularly in areas with low population densities (Wang et al., 2023). Additionally, MAUP also influences the comparison of accessibility derived from different population datasets as these datasets exhibit variations across different scales and regions. As a result, using boundaries with varying granularity introduces inherent biases.

In summary, this study demonstrates that 2SFCA results can be used to provide urban planning recommendations to the government and highlight directions for water facility improvements. Planning efforts should focus on integrating both the urban layout and the distribution of water facilities. Given the negative correlation between water quantity and accessibility distance (Cassivi et al., 2019, Martínez-Santos, 2017), these findings also contribute to research on water usage and water-related diseases. However, more accurate 2SFCA parameters, including kiosk capacity, travel thresholds, and population estimates, are needed to address the aforementioned limitations, requiring further detailed information. Future research could also incorporate field surveys to obtain records of queueing times at kiosks to improve these estimates. Additionally, analysing other shared water sources beyond kiosks would be beneficial for a more comprehensive understanding of water access patterns in SSA.

6.6 Conclusion

Shared water sources play a vital role in urban water supply systems in developing countries and have received attention in the pursuit of achieving SDG 6. In monitoring access to shared water sources, evaluations that consider spatial determinants are essential for providing policymakers with accurate and actionable information (Devi and Bostoen, 2009). Using Kisumu's kiosks as a case study, this study employed 2SFCA method to estimate accessibility, incorporating the distribution of population and kiosks while accounting for the decay effects of network distance on accessibility. Three gridded population datasets—WorldPop constrained, WorldPop unconstrained, and GHS-POP—were used to evaluate how different population disaggregation methods influence accessibility estimates and to ensure robust conclusions. The outputs were further analysed in relation to road centrality metrics to explore the impact of the urban environment on accessibility. Findings reveal that several slums within the study area, such as Manyatta and Migosi, achieved relatively high accessibility scores. This outcome is consistent across all three population datasets. The correlation analysis between accessibility and road centrality highlights two key conditions for improving water access for slum residents: a dense distribution of kiosks and strong transportation connectivity between the community and other parts of the city. Thus, improving accessibility in slums requires not only increasing kiosks but also enhancing the development of urban arterial roads. This study contributes to a deeper understanding of urban water heterogeneity and social inequality and provides valuable insights for advancing the goals of SDG 6.

Chapter 7 Discussion

The rapid urbanisation of Sub-Saharan Africa (SSA) has given rise to significant challenges in water supply and governance. Water access, encompassing both the provision of water and its proximity to residents, has been highlighted as a critical issue in the Sustainable Development Goals (SDGs) agenda. This study has examined these challenges across three key chapters.

In Chapters 4 and 5, an assessment of the overall and regional structural characteristics of pipeline networks provided insights into how urban development in SSA influences the configuration of water infrastructure networks. Notably, pipeline networks within slums exhibit significant structural differences compared to other urban areas. Chapter 6 further explored the demand and spatial distribution of shared water sources, revealing a strong alignment with the spatial organisation of urban layouts.

Building on these analyses, the following sections discuss how these findings contribute to a deeper **understanding of urbanisation, informality, and the heterogeneity of urban water services in SSA**, thereby supporting the overarching research objective of advancing progress towards the SDGs (Section 7.1). This discussion outlines the broader implications of the study for research and practice in water service provision (Section 7.2). Finally, key considerations regarding the limitations and transferability of the findings, as well as future research directions informed by this study, are explored in Sections 7.3–7.6, with a particular focus on their potential to enhance future practice.

7.1 Evaluation of thesis objectives

Objective 1: Assess the extent of topological and geometrical commonalities road and water/wastewater networks in SSA cities.

Observations of urban infrastructure network topology, combined with insights into pipeline network operation and maintenance, suggest that urban layout shapes pipeline structure, thereby affecting water service delivery (Section 2.3). Accordingly, Chapter 4 aimed to evaluate the hypothesis that, given shared urbanisation trends in SSA, and the role of roads and pipeline networks in urban planning, it should be possible to identify correlations between road and pipeline networks, with these relationships being influenced by the presence of informal settlements and other aspects of SSA urbanisation.

To achieve this, the study first developed a workflow for evaluating the connectivity and resilience of pipeline networks using graph theory metrics within two data-sparse case

study cities in SSA, followed by an analysis that quantified the differences in road networks between Kisumu and Kigali. These cities exhibit similarities in urban evolution, as described by Antos et al. (2016), where rapid spatial decay from the city centre is reflected in the graph properties of the road networks. The presence of both irregular and well-structured patterns influences the overall road network characteristics in the study areas. Given the differences in urban layout and informal settlement distribution between Kisumu and Kigali, their respective graph-theoretic properties display subtle variations—Kigali's road network exhibits lower connectivity efficiency but greater stability compared to that of Kisumu (see discussion in Chapter 4 for details).

Furthermore, the study **assessed the commonalities between road and pipeline networks across the case study SSA cities.** The findings in Chapter 4 indicate that infrastructure networks in both Kisumu and Kigali share common topological characteristics. The sewer network in Kisumu demonstrates lower structural stability than other networks, and its centrality measures show weaker correlations with those of the road network. By contrast, water pipeline networks consistently exhibit high structural similarity with road networks, with co-located roads and pipelines in both cities displaying comparable distributions of betweenness and closeness centrality. This suggests that pipelines in SSA follow common, scale-independent spatial configurations that are linked to urban morphology. Consequently, this finding not only identifies recognisable structural features in urban infrastructure networks, but also highlights the link between their spatial patterns and urbanisation—a connection that has so far lacked empirical support in research on cities of the Global South.

Objective 2: Assess spatial and topological differences in water networks between slums and planned urban areas.

This objective focused on a city-scale analysis of urban environmental factors. **Building on the previous chapter, Chapter 5 developed a workflow for evaluating the heterogeneity of pipeline networks using the Infomap community detection algorithm.** It examined the variability of water networks within the city by integrating local information to identify distinct network characteristics. Slum proliferation and delegated management model (DMM) were considered as potential influencing factors. Slum proliferation is prominent in SSA cities and is a possible explanation for the observed differences between urban infrastructure networks and road networks in the previous chapter. DMM, on the other hand, represents a prevalent community-led governance approach in SSA slums (see literature review 2.2.3). Both can be broadly regarded as products of informality in SSA urbanisation. The findings indicate that both factors correlated with pipeline network morphology. In addition to slums versus non-slums, pipelines managed under DMM also differ from other pipelines.

Objective 3: Quantify geographic accessibility to kiosk water using the two-step floating catchment area (2SFCA) model.

Given the widespread reliance on shared water sources as part of household water provisioning**, Chapter 6 investigated whether water point accessibility, like pipeline network topology, is shaped by urban form. Specifically, the study developed a workflow to evaluate geographic water accessibility using the two-step floating catchment area (2SFCA) method and provided recommendations for its application to water services. It also examined how different population attribution methods, accounting for the Modifiable Areal Unit Problem (MAUP), perform in 2SFCA analyses within SSA cities and correlates 2SFCA results with road network centrality.** The findings reveal that water point accessibility exhibits spatial heterogeneity within the study area, affected by two primary factors: the distribution of water sources and the spatial configuration of roads or overall accessibility. Moreover, access maps derived from constrained population datasets align more closely with the urban layout as indicated by the road network.

7.2 Research contributions

The primary contributions of this study are twofold. First, there are methodological contributions (Section 7.2.1): the application of graph-theoretic approaches and the 2SFCA method remains either unexplored or only minimally utilised in water research in LMICs. Thus, a key contribution of this study is the establishment of workflows for applying these methods in combination with urban planning data to characterise water distribution patterns in Global South cities. Following the sequence of methods employed within the overall research workflow, Section 7.2.1 presents the methodological contributions arising from the applications of topological correction, network generalisation, and community detection (Section 7.2.1.1), followed by the application of 2SFCA method to assess communal water point access (Section 7.2.1.2).

Second, through the application of these methods, this study identified key spatial characteristics of water service provision in SSA. These findings not only address existing gaps in empirical observations of urban water systems in SSA but also contribute to a more nuanced understanding of urbanisation theories. Section 7.2.2 first summarises the study’s key findings before reviewing gaps in existing theories on Southern urbanism and discussing how these results help bridge those gaps.

7.2.1 Methodological contributions

7.2.1.1 Application of network analysis methods to water pipeline infrastructure

In Chapters 4 and 5, this PhD study employed graph methods to study urban networks. As discussed in Chapter 4, graph-based analysis approaches to infrastructure often vary in their modelling methods (Marshall et al., 2018, Pueyo et al., 2019, Zhou et al., 2010). **Existing research on infrastructure network modelling has largely focused on road networks. In contrast, approaches for processing pipeline networks for graph-based analysis remain underexplored, while existing analyses of pipeline networks using graph theory are largely confined to case studies in** high-income countries**.** This study established a modelling and analysis workflow tailored to SSA pipeline network layers. During the data preprocessing stage, an evaluation and topological correction of the collected pipeline and road network data were conducted. Depending on the specific analytical objectives, the networks were modelled as dual and primal graphs in Chapters 4 and 5, respectively, and were standardised through graph smoothing to ensure consistency in their representation.

1. A topological correction workflow for water network layers in low-resource settings

The scarcity of graph analysis of water pipeline networks is not only constrained by the limited availability of data in LMICs (Yu et al., 2024), but also by issues of data quality. This not only includes the widespread problem of missing records, as exemplified in leakage analyses (Kirstein et al., 2019, Wu et al., 2024), but also frequent spatial and topological errors arising from infrastructure network management practices (Khaleghian and Shan, 2023, Solomakhina et al., 2016). In this PhD research, severe topological errors were likewise observed in pipeline datasets across both SSA case study regions. As a result, although topological approaches have been proposed as a simplified alternative to hydraulic modelling (Santonastaso et al., 2021), such analysis often remains unfeasible due to the lack of coherent topology in available data.

Topological correction is rarely addressed in the structural analysis of water pipeline networks, as hydrodynamic simulations typically rely on well-structured, real or simulated network data that do not require topological adjustments. As noted in Chapter 4, the issue of topological correction has been discussed only in a limited number of studies, primarily in the context of road network preprocessing—for instance, the *Shp2graph* package by Lu et al. (2018) and the *OSMnx* package introduced by Boeing (2017). **However, existing correction functions are not suitable for pipeline networks. Therefore, this study develops an approach for topological correction specifically tailored to pipeline networks.**

Addressing topological errors can be regarded as a form of network reconstruction, as precise knowledge of the actual pipeline distribution or the causes of errors remains unknown, similar to the challenges posed by missing data (Chittor Sundaram et al., 2020). Drawing on experience with urban network patterns, **the PhD study proposed a set of topological correction rules in Chapter 4**, such as using angles and distances to differentiate between topological errors and original/correct network structures. These correction rules could serve as a methodological reference for future studies on the topology of piped infrastructure in LMICs, facilitating the correction of raw datasets, improving data usability, and enabling the undertaking of additional case studies.

2. Water network generalisation

Before analysis, the representation of pipelines was addressed. As noted earlier, variations in data quality and mapping standards can lead to inconsistencies in analysis. Existing network generation techniques were all developed in relation to specific application contexts (Blagus et al., 2014, Cheng and Scherpen, 2021, Dias et al., 2018, Goyal et al., 2021, Maschler and Savic, 1999, Zhou et al., 2010), such as visual clarity or computational efficiency. Given that the networks in the study areas are relatively small, this thesis is less concerned with reducing network size for computational efficiency. Instead, the study focuses on ensuring consistency in network representation. In this study, the network was modelled as a simple graph, as is common in hydraulic studies (Goyal et al., 2021, Mah and Shacham, 1978, Momeni et al., 2023). Based on the generalisation principles, two main approaches were adopted in the PhD study. First, **during the identification process, as many elements as possible were retained in the graph**. This decision was made given that reducing the number of elements derived from the original network may lead to the loss of spatial properties, thereby hindering spatially informed analyses—such as the co-location analysis conducted in Chapter 4—from capturing certain spatial relationships. Subsequently, **a simplification process was conducted by removing pseudo-nodes that do not contribute to the network structure**, which are most likely introduced due to issues in network mapping. This approach ensures that graph-theoretical metrics derived from different networks maintain a consistent mathematical meaning. The simplification method was further developed in Chapter 5, where the extraction of pipeline circuits was largely guided by research on network simplification (Pung et al., 2022). This process is also related to the identification of tree structures and convex hull extraction in network analysis (Deuerlein, 2008, Šubelj, 2018). As a result of these steps, the networks are free from errors that may stem from data quality or capture issues. The resulting graphs preserve essential network characteristics—such as those relevant to hydraulic function—while exhibiting a consistent mathematical representation. This enables broader comparative and structural analyses. For example, in Chapters 4 and 5, the two graphs were analysed in both primal and dual forms, using nodes and edges, respectively, as units of analysis.

3. Applications of community detection

As shown in Table 2.5 of the literature review (Section 2.4.3), the application of community detection algorithms in infrastructure networks remains limited. Within existing studies on water networks, such algorithms have primarily been employed for the identification of District Metering Areas (DMAs), but not for examining the inter-relationship between urban form and water infrastructure. A few exceptions, such as Bramson (2022) and Law et al. (2019), have extended the range of applications by associating community structures within road networks with urban neighbourhoods. This approach inspired the PhD study, which adopts community detection as the principal quantitative method for identifying environmental factors associated with pipeline structure. **For this purpose, the research implemented community detection algorithms, complemented by interpolation, structural validation, and partition similarity assessments (Chapter 5). These approaches have not been previously applied to water network analysis, nor have they been explored from the perspective of water–urban morphology/governance interactions.**

This research introduces a combined approach that integrates multiple imputation with an interview-based vulnerability framework. The sequential imputation of missing values (IMPSEQ), a multiple imputation method that iteratively estimates missing entries, has performed strongly in water distribution systems (Kabir et al., 2020, Osman et al., 2018). This study also employed the pipeline deterioration factor scale developed by Marzouk et al. (2015). **The integration of these two methods effectively addressed the weighting requirements for community detection in this study and may offer a generalisable solution for handling missing vulnerability-related attributes in a wide range of pipeline networks.**

Following the implementation of community detection, this study incorporated a structural validation step to filter Infomap's multilayer outputs. Unlike validations that aim to evaluate the performance of community detection algorithms—typically by comparing detected communities with benchmark partitions that exhibit known community structures (Javed et al., 2018)—this study treated community detection outputs as a form of network attribute to be interpreted in relation to environmental variables. In such contexts, no ground truth exists against which results can be directly assessed. Therefore, Community Structure Validation (CSV) was employed to evaluate the structural strength of the resulting partitions. **This approach has not been widely applied.** The application of CSV demonstrated its ability to evaluate the overall structural validity of partitions across spatial scales. The results indicate that the most structurally meaningful layers in both networks are concentrated within the first three partition levels, a finding that corroborates the patterns observed in the subsequent partition similarity analyses. However, it was also observed that, due to its inherently conservative definition of community structure, the method tends to disregard smaller-scale communities at higher hierarchical levels, retaining only the larger ones and thereby reducing the size of the input partition. **Thus, the methodological contribution of this application is twofold: the incorporation of CSV into a workflow to enhance community detection for infrastructure networks, alongside recommendations on interpreting CSV outputs across hierarchical levels.**

To assess the correlation between the output partitions and environmental and governance characteristics, this study employed Adjusted Rand Index (ARI) and Adjusted Mutual Information (AMI), two metrics commonly used to measure the distance between partitions. Specifically, slum coverage and DMM information were combined to generate environmental partitions. This approach demonstrates a transferable method for aligning urban infrastructure networks with spatially distributed environmental data, with high potential for adaptation to a range of socio-spatial indicators beyond slum coverage.

In summary, Chapter 5 consolidates methods originally developed in other domains and contributes a coherent workflow for analysing the relationship between network structure and urban morphology or governance. **While none of the individual techniques are novel in themselves, their integrated application to the analysis of water infrastructure networks, particularly in the context of LMICs, is original and accompanied by detailed methodological guidance.** Furthermore, the proposed framework is readily adaptable, offering a flexible analytical structure that can be extended beyond water infrastructure or SSA cities.

7.2.1.2 Development of recommendations for applying 2SFCA models to water services

In the analysis of water point accessibility, the 2SFCA model was employed. As noted in Section 6.5, this accessibility model has not been applied in urban water supply studies, and it has only been used by a single rural study for assessing water point access (Mahuve and Tarimo, 2022). The general effectiveness of the 2SFCA approach is known to be influenced by several factors, particularly parameter settings such as threshold distances and distance decay functions (see Section 2.5.2). Previous studies have highlighted the importance of context-specific travel scenarios in ensuring the model's ability to capture realistic accessibility patterns (Chen and Jia, 2019, Wan et al., 2012). Accordingly, parameter calibration is essential when applying the 2SFCA model to urban contexts, where travel behaviours and spatial configurations differ from other settings. In addition, as noted in Chapter 6, the impact of population data on accessibility outcomes remains insufficiently understood. Prior studies have shown that different modelling approaches and covariates can lead to substantial variations across population datasets for the same area, which in turn affect accessibility estimations. Moreover, due to the scale and zoning effects of the MAUP, differences between population products in representing demand distribution may be amplified or mitigated by aggregation practices, yet this interaction remains underexplored.

To tailor the 2SFCA model for analysing water point access in urban SSA settings, this study refined both the parameterisation and population modelling components. Drawing on field-specific literature and case-specific observations, key parameters, such as travel mode, catchment threshold, and distance decay, were adapted to reflect realistic water-fetching practices in Kisumu. To simplify the complex water access landscape in Kisumu, this study excluded piped connections and other unrecorded communal sources, thereby focusing the analysis on one communal improved water source, namely water kiosks. Therefore, the study first estimated the population with access to piped water by multiplying domestic water meter density with mean household sizes per enumeration areas (EAs). These populations were excluded from the analysis, as their water needs were assumed to be met by piped connections. The remaining population was aggregated by EA and represented using population-weighted centroids to model demand for shared water points. Walking was adopted as the assumed travel mode, consistent with the findings of Crow et al. (2013) and Kim et al. (2020) on water collection in Kenya. **A 30-minute one-way catchment threshold was selected to account for the spatial distribution of water points and to avoid underrepresenting accessible sources.** Given that EAs rather than households were used as the population aggregation unit, the 15-minute threshold recommended by the JMP was deemed unsuitable in this context. The Gaussian distance decay function was introduced, and the Huff model was incorporated to weight accessibility scores based on proximity, capturing both the reduced likelihood of distant water collection and the impact of urban clustering of water points on destination choice. **These access parameterisation choices, along with the rationale behind them, may inform future water access research in urban contexts across SSA.**

To compare how different population datasets perform in accessibility modelling in SSA urban contexts, this study included three datasets with substantial differences in modelling approaches, covariate selection, and methodological assumptions: WorldPop Global Constrained, WorldPop Global Unconstrained, and the GHS-POP. Both the WorldPop Constrained and GHS-POP datasets incorporate built settlement extents as constraints in their population distribution processes. To reduce the influence of the scale effect of the MAUP, EAs were used as the unit of aggregation when computing population-weighted centroids. **Despite this, notable variation in accessibility patterns remained across the datasets, indicating that differences in population allocation still affect spatial estimates of access.** Among the three, the two constrained datasets produced more consistent results with each other, while the unconstrained WorldPop dataset diverged more substantially. Notably, the WorldPop Constrained dataset exhibited the highest spatial alignment with road centrality patterns, potentially due to the inclusion of detailed building footprints in its modelling workflow, which constitutes a further theoretical advantage of constrained models in capturing the influence of urban form on accessibility. **Given the 2SFCA model's sensitivity to spatial distribution and distance, this alignment indicates that the WorldPop Constrained dataset is the most reliable population data source for water access analysis in this setting.**

7.2.2 Evidential and theoretical contributions

In SSA, the advancement of both SDG 6 (clean water and sanitation) and SDG 11 (sustainable cities and communities) faces pressing challenges, notably reflected in the rapid expansion of slums and the difficulties in providing adequate water services (UN-Habitat, 2023b, UN-Habitat, 2023c, UN, 2023a). The SDGs proposed by the United Nations exhibit multiple interlinkages, both in their conceptual definitions (Coopman et al., 2016, Le Blanc, 2015) and in regional implementation practices (Griggs et al., 2017, Herrera, 2019), whereby actions targeting a particular indicator inevitably influence the achievement of other goals, either positively or negatively (Abson et al., 2017, Nilsson et al., 2016). Specifically, SDG 6 and SDG 11 have been recognized to exhibit both synergies and trade-offs (see Section 2.1.3). In practice, SSA's water and slum challenges are closely tied to rapid urban population growth, stemming from a fundamental mismatch between service provision and urban demand. This leads to both conceptual and spatial intersections between slums and unmet water needs (Adams et al., 2019, Dos Santos et al., 2017). Regarding piped water services, earlier studies suggest that slum expansion contributes to intra-urban disparities in service provision (Lagerberg, 2016, Mapunda et al., 2018, Shushu et al., 2021). Moreover, due to the path-dependent nature of network evolution, pipeline infrastructure within slum areas often exhibits distinct structural features, which likely affect service performance, as summarized in Section 2.3.2. Meanwhile, water access studies on both water source types and accessibility exhibit spatial heterogeneity (Deshpande et al., 2020, Dongzagla et al., 2022, Tetteh et al., 2022). However, despite these insights, **systematic and spatially explicit evidence on the coupled relationship between slums and water infrastructure remains scarce**. In particular, few studies quantitatively explore the mechanisms through which piped and non-piped water access are shaped by urban spatial structure. This evidential gap poses two key challenges. First, from a knowledge perspective, the absence of empirical data hinders a systemic understanding of how urban planning affects both marginalised urban populations and the infrastructure systems that serve them. Second, from a policy perspective, interventions—such as slum upgrading or water utility reform—often lack a nuanced grasp of this spatial interaction, impeding the development of targeted, context-sensitive strategies. A case in point is the DMM implemented in Kisumu, which was explicitly designed to address water supply issues within slums by replacing the chaotic 'spaghetti' network with a more structured one (World Bank, 2009). While such initiatives directly respond to the challenges outlined above, there is to date no empirical evidence on their effectiveness, highlighting the need for further empirical evaluation (Nzengya, 2015). This study addressed this gap by applying an integrated analytical framework that combines urban morphology and water infrastructure analysis, **revealing that pipeline topology is shaped by broader urban environments, including both physical form and governance strategies such as the DMM. This approach enabled the empirical identification of spatial interactions and trade-offs between SDG 6 and SDG 11 at the intra-urban scale**.

The overarching approach of this study is to examine the performance of road networks, water and sanitation infrastructure, with particular attention to how these associations vary across slum and non-slum areas, as well as under different management regimes. In sub-Saharan Africa, urban water provision encompasses both piped and non-piped sources. This study focuses exclusively on two improved water sources, pipelines and water kiosks, which are representative examples of piped and non-piped water supply, respectively. The morphology and management of pipeline networks influence maintenance and water supply stability (see Section 2.3.2). The research findings indicate that, despite differences in scale and national context, infrastructure networks in SSA cities exhibit remarkable commonalities in their structural configuration, with connectivity and robustness being closely linked to urban morphological characteristics (see Chapter 4). Notably, the D-measure and co-location analysis of road and water networks—particularly the linear relationship observed in co-located road and pipeline closeness centrality—highlights two key points: (1) urban morphology has a strong relationship with infrastructure networks, and (2) the road and water networks of two case study SSA cities share common morphological characteristics. The significance of this study lies in highlighting the influence of urbanisation in the Global South on the configuration of water systems. Crucially, the influence is independent of specific conditions such as city size or geographical location, as evidenced by structural similarities observed across comparable types of infrastructure in different cities. The findings provide a foundation for analysing the relationship between slums and water network configurations. Previous studies, despite their efforts to compare water infrastructure networks across cities using graph-theoretical approaches, have generally failed to reach clear conclusions due to the absence of quantitative measurements (Abdel-Mottaleb and Zhang, 2020, Lorenz et al., 2021, Mair et al., 2017). Further analysis (Chapter 5) examined the spatial configuration of pipeline layouts within the study cities. The results demonstrate that the heterogeneity of network structures within the cities, where both government intervention and the spatial configuration of slums are related to the variations in pipeline network structures at the intra-urban level. These findings are further supported by graph metrics and supplementary contextual information. In particular, pipelines in Kisumu managed under DMM show distinct characteristics in community detection, indicating that they tend to form discrete, self-contained subsystems compared to pipelines in other areas.

Regarding water kiosks, existing studies, such as Adams (2018b) and Boakye-Ansah et al. (2022) listed in Chapter 6, have largely relied on small-scale surveys and interviews. While valuable for uncovering specific local details and consumer experiences, these studies do not offer broader spatial or quantitative insights into access patterns. As a result, they provide limited guidance for large-scale urban planning or policy formulation. Chapter 6 reveals that EAs with high accessibility values are concentrated in certain slums where kiosks are most densely distributed. This variation between slums in kiosk availability also corresponds with earlier findings of prolonged queuing times in other slum areas with lower kiosk density (Sima et al., 2013). Accessibility declines outward from these areas. Meanwhile, for water points like kiosks, the relative spatial distribution of these points in relation to the population directly affects their accessibility and, consequently, water consumption (Boakye-Ansah et al., 2022, Cassivi et al., 2019, Devi and Bostoen, 2009). This pattern can also be linked to urban morphology, as it overlaps with the KDE and IDW interpolation results of road network centralities, which serve as indicators of major urban road density. This finding is consistent with the commonly observed heterogeneity of water provision in SSA cities (Deshpande et al., 2020, Dongzagla et al., 2022, Tetteh et al., 2022), but offers more data-driven evidence and identifies the spatial factors that may most effectively enhance access.

Our study provides a robust quantitative framework to address these gaps by demonstrating that the heterogeneity of both pipeline and water point services can be recognised by a common environmental factor, urban morphology. Urban morphology is determined by the spatial organisation of city elements—primarily plots, buildings, and streets (Araújo de Oliveira, 2022b, Scheer, 2015), among which streets are commonly used as a basis for analysis (Araújo de Oliveira, 2022a, Wang and Gu, 2023). In the PhD study, road networks were used as a reference in both Chapters 4 and 6, while Chapter 5 employed the Million Neighborhoods Map, a dataset that also utilises urban morphology to identify slums. Urban morphology influences pipeline placement by determining spatial divisions, leading to structural variations within slums. Due to differences in construction guidelines between streets and pipeline networks, a systematic centrality difference between the two was observed. As a result, the PhD study observed similar infrastructure metrics across both cities in Section 4.4. This was further validated through community detection and partition similarity assessments, which confirmed that pipeline structures within slums exhibit distinct network characteristics. Additionally, in Kisumu, community detection revealed that the DMM had a notable impact on pipeline topology.

For water points, the study's findings align with intuitive expectations: urban morphology and the spatial distribution of water sources are both associated with accessibility. In Kisumu, the clustering of high accessibility scores overlapped with densely distributed kiosk areas and regions with high road centrality. These findings provide further evidence that SSA cities exhibit identifiable structural patterns that influence both infrastructure provision and residents' daily lives. Furthermore, pipeline network arrangements and water interventions (including both DMM pipelines and kiosks, the latter of which in Kisumu is also linked to the DMM) represent an attempt to formalise informality. The observations indicate that formal and informal networks exhibit distinct morphological characteristics. Neoliberal water governance policies have made measurable progress, aligning with Harris (2020), who argues that while neoliberal reforms are often perceived as a withdrawal of formal governance and require careful scrutiny, they have nonetheless objectively improved essential services.

Overall, these findings provide strong evidence to support the developing theories of Southern urbanisation. Since the 1960s, scholars have proposed varying interpretations of informality, offering insights into how it should be conceptualized and positioned within the urban context. A systematic review of urban studies indicates that the focus on cities in the Global South is increasing (de Castro Mazarro, 2023). This shift has been accompanied by changes in both policy and research perspectives, moving from the experiences of northern cities to the realities of southern cities, as well as a transition from urban renewal strategies to subaltern urban governance (Brenner and Schmid, 2015, Sheppard et al., 2013). Scholarship has advanced alternative theories of urbanisation that seek to incorporate Southern urban experiences into a more inclusive theoretical framework, while also exploring the underlying drivers and conceptualisation of such practices (Parida and Agrawal, 2023). This theoretical trend in Global South urban studies, along with the associated developments in slum and water policies, is described in the literature review (Section 2.2). In this context, merely increasing the number of studies on cities in the Global South does not necessarily enhance understanding of their development; rather, it risks reinforcing fragmented perspectives on specific cities. **Therefore, research on Southern cities, particularly on informal slums, is essential for deepening the understanding of Southern urbanisation.**

Studies of Southern urbanisation primarily focus on how to approach the existence of slums. The subaltern perspective recognizes the value of the activities of slum residents, rather than viewing them merely as objects to be excluded. From another angle, the resistance of slum residents to slum clearance activities can also be seen as an expression of subaltern power (Fix and Arantes, 2022). This is the reason behind slum upgrading policies discussed in the literature review and the introduction of the DMM / public-private partnership (PPP) model for improving slums. The current research demonstrates that such intervention policies, distinct from full public actions, can effectively change service provision in neighbourhoods. **The deeper conceptual issue that the theories address is how to identify the poverty entities that require governance.** As Roy (2014) pointed out, since the SDGs focus on eliminating urban inequality and poverty, recognizing and understanding the poverty entities that need improvement in cities becomes a key issue. This involves determining which rights or conditions should be seen as essential for urban residents. As noted in the literature review 2.2.1, different countries define slums in various ways, reflecting their understanding of urban inequality. Alizadeh and Prasad (2024) advocate for *The right to the city*, which views the rights of excluded urban residents as a basis for identifying urban inequality. **Following this perspective, in the PhD study, whether measuring network structures or water kiosk access, the essence is to treat access to basic services as a universal right for urban residents, and, based on this, quantitatively assess the risks of being excluded from these rights.** Through comparison with areas commonly regarded as slums, valid similarities in their distribution were found, **indicating that service provision can be viewed as a key indicator of urban inequality.** This suggests that the observed inequalities reflect broader, shared patterns of urban development imbalances. Therefore, the evidence contribution, which concerns inequalities in access to water services, is not limited to the context of slum-free cities or Southern urbanisation.

7.3 Research limitations and uncertainties

The limitations of the study can be divided into three categories. The first category consists of the shortcomings arising from practical implementation, which can be improved with more accurate information. The second category refers to the inherent limitations of the research methods and approach; these limitations are intrinsic to the methods used and can only be addressed by employing alternative approaches. The third category stems from external factors, such as policy frameworks and socio-economic conditions, which impose contextual constraints on the research.

Limitations that can be addressed through acquiring more robust information: First, the lack of access to temporal data restricted this analysis to a cross-sectional study. In such cases, causal inference is generally inappropriate without additional supporting information (Kesmodel, 2018). As a result, this study was limited in its ability to explore temporal relationships between urban environmental change, policy interventions, and the evolution of pipeline networks. Specifically, the study was unable to determine whether the observed relationships hold consistently over time or confirm a causal link between urban morphology and networks.

In the 2SFCA analysis, I simplified the water access situation within the study area by applying the same catchment size to both the population centroids of peripheral and internal EAs. The EA boundaries for Kisumu were derived from the 2009 Kenya Census rather than the more recent 2019 Census, due to data availability. Moreover, this study did not account for the diversity of water sources within the study area. In this study, kiosk demand was estimated by subtracting the population using piped water from the total population. However, in many developing contexts, including Kisumu, households often rely on multiple water sources for different purposes (Daly et al., 2021, Okotto et al., 2015). Due to the lack of detailed data on how water use is allocated across sources in Kisumu, this study adopts a simplified estimation approach that does not fully reflect such complexities. Moreover, the travel scenarios and distance decay parameters were set based solely on existing literature rather than field surveys in Kisumu. Therefore, the access analysis results may differ from actual water collection conditions. These limitations, similar to those in graph-based analysis, could be improved with further investigations and the collection of field data.

Inherent uncertainties stemming from modelling assumptions and data limitations: Further challenges arise from the application of graph theory and the 2SFCA method. The interpretation of topological attributes is influenced by other similar studies. On one hand, there is no established standard for how to apply graph measures to infrastructure networks, which introduces methodological variations in network modelling and representation, such as whether to use primal or dual mapping, and how to correct and simplify network topology. This results in potential differences in the interpretation or significance of the graph metrics employed. For example, Hwang and Lansey (2017) pointed out that removing or retaining pseudo-nodes within the same network can result in different meshedness values. This implies that differences in modelling assumptions can render network indicator values non-comparable across studies. This issue is also related to the availability of data. Many studies of water network topology rely on synthetic networks, which do not always reflect the characteristics of real-world networks (Momeni et al., 2023, Paez and Filion, 2017, Yu et al., 2024), thus undermining the comparability of the study's findings with other studies. Another limitation of graph theory analysis is that, although network topology is closely linked to the hydraulic performance of pipeline networks, it should only be viewed as a preliminary basis for network resilience analysis. As Yazdani and Jeffrey (2011) point out, topological analysis alone provides an incomplete picture of network resilience, as financial and operational management, geographical context and the urban space heavily influence the structure of the network.

In addition, several inherent uncertainties in the accessibility analysis cannot be resolved through the acquisition of better data or the adoption of alternative methods. First, gridded population datasets are inherently predictive in nature and rely on modelling assumptions that introduce uncertainty. Although this study compared multiple population datasets through assessing their alignment with road centrality patterns, the results can only indicate relative performance. Second, household-level data on population distribution and water access would offer higher accuracy but are often unavailable due to geoprivacy concerns and the limited spatial coverage of detailed surveys. Third, despite being a critical component of actual water access, unimproved water sources are often informal, transient, and difficult to monitor, making them challenging to incorporate into spatial models with consistency and reliability. As a result, analyses of water access are inevitably subject to uncertainty, since actual service levels can only be approximated rather than precisely measured.

Limitations arising from socio-economic and governance factors: Moreover, the interpretability of this study is subject to variability due to governance and socio-environmental factors. A notable limitation lies in the definition of slums in Chapter 5. Extracting urban structure and community information is challenging due to the lack of widely accepted definitions for communities (Harris, 2020). This is particularly true for slums (Kuffer et al., 2016). In this study, slum boundaries were derived from the Million Neighborhoods Map and supplemented using UN-Habitat slum maps. Although the Million Neighborhoods Map also accounts for population distribution, it fundamentally relies on the analysis of road network layouts, meaning its results may still differ from studies employing alternative slum identification methods. Another greater source of uncertainty arises from the fact that, due to political and social factors, slums exhibit a wide variety of forms and compositions (Smit et al., 2017, Taubenböck et al., 2018), and definitions differ across countries and even between cities. Consequently, using a broad classification of slums or informal settlements may result in variability in analyses of slums in relation to urban pipelines, depending on the criteria applied. This limitation cannot be fully resolved until a unified framework for defining community and slum boundaries is established. Furthermore, in the community validation section of Chapter 5, other environmental factors potentially affecting pipeline distribution may have been overlooked. Variations between slums in terms of population, infrastructure conditions, and income levels, as well as differences in terrain and the priorities of government strategic planning, may exert additional influence on the configuration of pipeline networks within communities.

Similarly, the study also reveals that urban morphology plays a role in water access. However, in Kisumu, water kiosks display a notably clustering distribution. Considering that the establishment of water kiosks is often shaped by power dynamics (Velzeboer et al., 2018), this dense distribution suggests that kiosk placement is likely influenced by governmental decisions or environmental factors. It may also reflect heterogeneity in community demand for kiosk services. Therefore, an analysis focusing solely on access cannot fully capture the relationship between supply and demand, and the layout of the road network represents only one of the factors influencing access.

7.4 Transferability and scalability of the research

Although the setting of parameters requires local knowledge, the analysis approaches employed in this study are not dependent on the characteristics of specific cities and can be applied to other cities, including graph similarity measures, community detection algorithms, co-location analysis, and the 2SFCA model. Graph methods treat networks as the mathematical representations, preserving only the connectivity between nodes, while other attributes are incorporated as alternative weights. Moreover, network metrics are designed to account for the varying numbers of nodes and edges within the network. The study's findings also show that graph-based analyses exhibit characteristics that are scale-independent, especially in the case of co-location analysis. While this study adopted a specific community detection algorithm, Infomap, it is important to note that no single algorithm is universally optimal. Due to the absence of a definitive ground truth and the context-dependent nature of algorithm performance—summarised in the *no ground truth* and *no free lunch* theorems (Fortunato and Hric, 2016, McCarthy et al., 2019, Peel et al., 2017), researchers seeking to apply this workflow to other networks should be cautious about how algorithm choice may affect results.

The 2SFCA model, as a place-based accessibility assessment approach, is fundamentally an application of Tobler's First Law of Geography (Tobler, 1970), evaluating accessibility by measuring the supply-to-demand ratio and the spatial proximity between them. Therefore, the workflow developed can be applied where information on population, water points, and travel scenarios is available. Among these, population datasets are relatively well-developed, with many offering global coverage—such as the HRSL, GHS-POP, LandScan, and WorldPop global layers listed in Section 6.3.1, as well as the forthcoming WorldPop Global 2. Considering the differences observed in the performance of population datasets in 2SFCA analysis and their timeliness, some additional population indicators can also be used to assist in population estimation (Tan et al., 2021). However, population datasets differ in accuracy depending on modelling approaches, input covariates, and data availability—particularly in slums, where covariates such as building footprints may be insufficient or outdated (Hanberry, 2022, Palacios-Lopez et al., 2019, Thomson et al., 2021). As this study has shown that constrained datasets performed better within slum areas of the study region, we recommend careful evaluation of population layers when applying the 2SFCA workflow to other contexts. On the other hand, studies have already established global and national-level travel scenario models. For example, Macharia et al. (2023) calculated the geographic accessibility of public primary schools in Kenya, using Accessmod software to construct travel paths, a method also employed in the current study. Weiss et al. (2018) created global city accessibility maps, also employing the friction surface modelling method. Watmough et al. (2022) constructed travel time maps to the nearest health facility across Uganda, Tanzania, Zimbabwe, and Mozambique. **These examples highlight the increasing availability of large-scale population and travel scenario data, which supports the application of the 2SFCA model.**

However, the application of the 2SFCA model requires a detailed and georeferenced inventory of communal water points to construct travel paths between populations and specific sources. While several countries—such as Sierra Leone, Liberia, Uganda, and Tanzania—have national-level water point databases, studies have highlighted persistent quality issues, including procedural, observational, and conceptual errors (Foster, 2013, Yu et al., 2017). These issues, arising from both local contexts and general limitations in water point mapping, can introduce errors when combined with population estimates (Verplanke and Georgiadou, 2017, Yu et al., 2017). The study's data audit also shows that the generation of population data products relies on census data, which means that when applying the 2SFCA analysis in other areas, there may be discrepancies between the model outputs and actual supply-demand conditions. **Nevertheless, the method remains promising in contexts where the reasonably complete and up-to-date inventories of water points are increasingly available through systematic mapping campaigns and the digitization of utility records.** In addition, platforms such as the Water Point Data Exchange (WPdx), the International Benchmarking Network for Water and Sanitation (IBNET), SIASAR, and the JMP offer valuable reference sources to support broader application and calibration of the model at city or national scales.

Our focus is on addressing urban inequalities, particularly in spatial and water services. Apart from the availability of pipeline data, the only criterion for selecting study areas was that they should face challenges of slum expansion and water access. This criterion ensures the representativeness of the study and allows for transferability of the results across various SSA cities. Therefore, the two selected cities differ significantly in other aspects, such as city scale, spatial layout, infrastructure development history, and water policies (Section 3.2). Despite these differences, both cities still exhibit similar infrastructure patterns, as demonstrated by the findings in Chapters 4 and 5. This suggests that the observed relationships between urban form and infrastructure may also hold in other SSA cities with urban inequalities, regardless of their other conditions. Furthermore, as previously discussed, the development of urban infrastructure—whether in southern or northern cities—is constrained by similar factors. The study's focus on slums and water services in SSA cities fundamentally reflects a broader pattern of urban water distribution inequalities. **From the perspective of comparative theory, this research eventually contributes to the understanding of global urban development.**

7.5 Implications for urban water service delivery and planning

In this PhD study, the application of graph theory methods demonstrates significant potential for evaluating pipeline networks and informing pipeline planning, particularly in LMICs where data on pipelines is often scarce. However, it remains important to maintain comprehensive records of pipeline infrastructure to support accurate network performance assessments. As noted in the limitations, the absence of installation date records hinders the possibility of conducting longitudinal studies on pipeline networks. Therefore, utilities are encouraged to improve the management of pipeline data and maintain detailed records of installation dates. In addition, documenting other physical attributes such as pipe diameter and material is also recommended, as this information can contribute to more accurate weight estimations in network analysis.

As highlighted in the literature review (Section 2.2.3), the implementation of PPP projects, such as the DMM, has been subject to considerable criticism worldwide. Many of these projects have either failed (Ameyaw and Chan, 2015, Zhang and Tariq, 2020) or are perceived to have shown no significant performance improvements compared with publicly managed initiatives (Chenoweth and Bird, 2018). As a result, PPPs are often regarded as involving substantial risks. On the other hand, from the perspective of slum upgrading, existing research has demonstrated that the public-oriented nature and social engagement inherent in PPPs can bring significant benefits to upgrading projects (Brown-Luthango et al., 2017, Svensson et al., 2003). In this context, the extent to which PPPs can improve existing conditions becomes a crucial question. For water provision, given the crucial role of pipeline networks in determining water access and their long lifespan, which implies long-term impacts, it is also essential to examine their physical configuration. Evidence from Chapter 5 indicates that pipelines implemented under the DMM project exhibit distinct characteristics compared to other areas. Since one of the key motivations behind the DMM deployment is to improve the disordered "spaghetti" pipelines commonly found in slum communities (World Bank, 2009), **the results suggest that these interventions have successfully enhanced water distribution networks in slums**. Although the nature of this impact—whether positive or negative—remains unclear, it nevertheless provides an important indication that PPPs have the potential to continue or expand their role in improving water provision in slum areas through interventions in the pipeline network.

Furthermore, this suggests the feasibility of the Southern urban-style transformation towards sustainable water management advocated by (Herslund and Mguni, 2019).

This research also indicates that there are still certain shortcomings in the current placement of DMM endpoints. In the Kisumu area, there are two operational management models for water kiosks. One is managed by the DMM master operators, who collaborate closely with the utility. The other is operated directly by the utility and kiosk operators as separate actors, without the involvement of master operators. The DMM kiosks are mainly concentrated in Nyalenda, while the kiosks in Manyatta follow the latter model (two areas identified in Chapter 6) (Gilson et al., 2025). According to a survey by Nzengya (2018), it was observed that the number of people queuing at kiosks in Nyalenda is significantly higher than in Manyatta, despite the kiosks in Nyalenda offering lower prices. The study's observations support this phenomenon, as Nyalenda shows lower access levels in the accessibility analysis, while the areas around Manyatta have higher water access. The former pattern is likely primarily related to supply shortages and may also indicate affordability constraints where limited competition leads to higher prices (Section 6.5). Therefore, while DMM kiosks play a role in extending water access to underserved communities, **their limited distribution constrains broader impact. This finding points to the importance of expanding kiosk provision, particularly in areas with constrained access and limited inter-kiosk competition**. Policymakers may consider introducing spatially targeted support mechanisms, such as start-up financing, training programmes, and streamlined administrative procedures, to facilitate the entry of new kiosk operators and reduce barriers to service provision.

From an urban planning perspective, one of the major challenges faced by cities in the Global South is how to address the expansion of slums that accompanies rapid urbanisation. While in-situ slum upgrading has been implemented in many contexts (Bolton, 2020), research has also pointed out that some projects have achieved only limited improvements in residents' access to basic services(Edith et al., 2019, Patel, 2013). This places higher demands on infrastructure planning.
Adopting a morphological perspective, this study highlights several key considerations for planners:

1. **The results indicate a strong relationship between the pipeline network and the distribution of slums within the study area.** This suggests that the poor condition of infrastructure in slums is closely related to the spatial configuration of service facilities. Consequently, improvements in accessibility and performance should be grounded in modifications to the spatial layout of these facilities.
2. **Infrastructure upgrading should not only focus on the internal spatial planning of**

slums, but also take into account their interaction with facilities in adjacent areas, as well as their location within the wider urban infrastructure network — for example, their distance from main roads or trunk pipelines, and their connectivity with other communities. **3. The overlap between roads and pipelines indicates that slum upgrading plans need to consider how the allocation of public space influences the spatial distribution and interaction of multiple infrastructure networks.** Furthermore, due to the spatial misalignment between road and pipeline networks caused by their different core layouts and design principles, planners must also recognise that such interactions are constrained by local conditions, for example, the locations of reservoirs, treatment plants and trunk lines, or the road structure of well-developed areas such as CBDs or historically established neighbourhoods.

7.6 Future directions for research

Future research can address the gaps identified in this study and expand on its findings. As noted in the limitations section, the current study has methodological shortcomings. Future work can improve the framework by:

1. **Enhancing the identification of urban inequality.** In this study, the assessment of slum impacts is based on the Million Neighborhoods Map, but future research could explore alternative approaches to identifying slums (Kuffer et al., 2016, McCartney and Krishnamurthy, 2018, Smit et al., 2017). Whether for slum-related research or for applying this framework to other cities, it is essential to clarify what urban inequality means in the specific research context and how disadvantaged areas should be identified. This remains an important aspect for future research.
2. **Exploring the complex interactions between water points, population distribution, and urban morphology to refine 2SFCA assessments.** This is crucial for further developing accessibility analyses to improve water service provision.
3. **Leveraging alternative methods for measuring network similarity beyond the D-measure.** Neural network-based approaches have gained traction in urban morphology research. For instance, Kempinska and Murcio (2019) employed Variational Autoencoders (VAEs) to encode urban street networks into low-dimensional representations, enabling quantitative analysis of CNN without prior domain knowledge. Ma et al. (2024) utilized GCN-based models to predict the complexity and connectivity of street networks across cities. These methods can provide comparative insights and enable broader network analysis. Additionally, methods assessing node similarity within

networks, such as SimRank* (Yu et al., 2019a), can also help quantify regional characteristics within the network.

The robustness of the research can be further improved through enhanced data collection and analysis. As discussed in Section 7.2.1, the topological correction of the network essentially involves a reconstruction of the data based on prior knowledge. Collaborating with individuals who possess contextual expertise, such as engineers familiar with the local network, could improve the reliability of the corrected network and enhance its alignment with the real-world infrastructure. With more detailed temporal data, future studies could also analyse the evolution of pipeline networks and further depict the relationship between urban morphology and pipeline infrastructure. Additionally, as highlighted by Sambu (2016), water governance strategies in Africa have undergone significant changes over the past few decades. Therefore, a longer-term observation of pipeline networks would provide deeper insights into how policies have influenced the configuration of urban networks. Previous studies have made noteworthy attempts in this area. For instance, by examining graph metrics of Hong Kong's road network over time, Lan et al. (2022) demonstrated that urban networks exhibit a specific evolutionary trajectory. In the study of the changing graph properties of Paris's road network over two centuries, Kirkley et al. (2018) found that centrality indicators were insensitive to changes in the urban spatial layout. Similarly, Krueger et al. (2017) observed that, over several decades, water distribution and sanitary sewer networks in an Asian city displayed nearly stable graph characteristics that were independent of urban settings, suggesting that infrastructure evolution follows generic mechanisms. **Given the similar phenomena observed in the study, conducting such research in SSA regions would likely yield insights into evolutionary trajectories for infrastructure.** Furthermore, Sulem et al. (2024) proposed a method for detecting change-points by comparing the similarity of continuously changing graphs. **This suggests that the graph similarity analysis techniques employed in the study could also be further enhanced when applied to the analysis of temporal data.**

As highlighted in the limitations section, more robust 2SFCA estimation require records including kiosk capacity, travel thresholds, queueing times at kiosks, and population estimates for kiosk usage (or other water sources, depending on research objectives). Therefore, incorporating detailed field surveys to improve these estimates with detailed and up-to-date information would be valuable in future studies. This requires further communication and investigation with utilities and master operators within the study area to gather information about project implementation and kiosk operations (e.g., water pricing, service coverage, operating hours) and to understand their challenges. This should help refine the reliability of the model and provide more meaningful and realistic interpretations of water services.

Furthermore, with access to additional data, future studies could examine the implications of the observed topology and accessibility characteristics using the analytical framework developed in this study. As reviewed in Section 2.3.2, a range of studies has demonstrated that pipeline topology is associated with key aspects of water service, including leakages(Adraoui et al., 2024)， hydraulic characteristics(Torres et al., 2017, Yu et al., 2024)， energy consumption for water supply(Zhao et al., 2020)， and asset management(Godfrey and Howard, 2005a). Therefore, future research could apply the methodological framework developed in this study to identify structural heterogeneity in networks and examine whether such differences are reflected in operational outcomes. For example, by developing a geo-referenced database of pipeline breakage incidents, researchers could compare failure frequencies between informal and formal areas and assess whether these are associated with specific topological features or community structures. **Similarly, future studies could explore the relationships between 2SFCA-based accessibility and other attributes of water services. Several studies have investigated the affordability of kiosks and reported spatial disparities in water prices** (Adams, 2018b, Adusei et al., 2018, Boakye-Ansah et al., 2022, Nzengya, 2018, Ondigo et al., 2018, Tumwebaze et al., 2023). These price differences are often linked to variations in the density and spatial distribution of water points, which may reflect levels of market competition or monopoly. Notably, Nzengya (2018) reported spatial variations in water pricing in Kisumu, and these patterns appeared to correlate with the accessibility results derived in the PhD study, which likewise accounted for the spatial distribution of water sources. These findings suggest that the results of 2SFCA analysis may serve as a useful proxy for understanding affordability. If price data were available, it would be possible to examine such correlations more directly, similar to how this study related road network centrality to water point accessibility. This approach could also be extended to investigate broader dimensions of service delivery, including availability, affordability, and water quality.

The findings of this study also open several new directions. For example, integrating accessibility analysis with network analysis, as demonstrated in Kranioti et al. (2022)'s study, where accessibility between locations was used to construct network edges. Moreover, multilayer network research has increasingly focused on the interdependencies among urban infrastructure systems Building on this perspective, future research could advance efforts to conceptualise infrastructure as an interconnected system by adopting a multilayer network approach, thereby examining structural interactions or synchronisation across co-located networks such as water, electricity, gas, and roads (Boccaletti et al., 2014, Kivelä et al., 2014). Furthermore, this PhD study annotated network features using slum and DMM information， in order to examine their associations with network structure. This approach could be extended to other network or environmental factors. One possible direction could involve adopting the method proposed by Müller et al. (2020) to delineate slope-based zones, assigning networks within the same landslide risk area to a common community. By comparing the resulting partitions, future research could assess whether such factors influence network structure.

Chapter 8 Conclusion

Urbanisation in the Global South has resulted in diverse urban forms and water service delivery practices, exhibiting characteristics distinct from those of the Global North. Previous research, which predominantly relies on surveys and interviews, has largely failed to capture the broader-scale characteristics and influencing factors of water service provision. This study contributes to filling the gap in quantitative evidence by applying graph theory and accessibility measures to examine the spatial heterogeneity of water services in Sub-Saharan African (SSA) cities. Further, the relationships between water services and contextual factors such as water supply policies and community characteristics are explored. Chapters 4 and 5 employed graph methods to assess the performance of piped networks from two different SSA cities (Kisumu, Kenya; and Kigali, Rwanda). The findings revealed that the water pipe networks in the two cities exhibited remarkable similarities, with network attributes strongly associated with the commonalities observed in SSA urbanisation. Building on this, differences in urban planning strategies between the two cities were reflected in variations in network connectivity and robustness. The presence of slums contributed to the formation of distinct community structures within the piped networks of both cities, which were further differentiated by water interventions specifically targeting slums. The analysis of access to water kiosks addressed a significant research gap concerning the spatial distribution of off-premises water supply in SSA cities and potentially explained variations in kiosk water pricing across slums. The analysis revealed that, beyond supply and demand factors, the characteristics of shared water provision are also associated with road centrality. Furthermore, the study indicated that access estimates derived from population datasets that are constrained by built settlement extent exhibited a stronger correlation with road centrality, suggesting a closer alignment with urban spatial structures.

This research established an analytical workflow that applies methodologies developed in other fields to the study of water infrastructure, enabling a comprehensive and large-scale assessment of both piped and non-piped water services within the unique context of SSA. In doing so, the study contributes to bridging existing gaps in the literature on water practices in the region, while also shedding light on the interaction between SDG 6 and SDG 11 via the lens of water infrastructure. Building upon the findings of this study, future researchers and urban planners may gain deeper insights into the dynamics of urban development, thereby informing more equitable and sustainable urban planning practices and advancing towards the achievement of the SDGs.

List of References

Abass, A. S. & Kucukmehmetoglu, M. 2021. Transforming slums in Ghana: The urban regeneration approach. *Cities,* 116**,** 103284.

Abdel-Mottaleb, N. & Zhang, Q. 2020. Water Distribution–Transportation Interface Connectivity Responding to Urban Geospatial Morphology. *Journal of Infrastructure Systems,* 26**,** 04020025.

Abebe, G. & Hesselberg, J. 2015. Community participation and inner-city slum renewal: relocated people's perspectives on slum clearance and resettlement in Addis Ababa. *Development in Practice,* 25**,** 551-562.

Abson, D. J., Fischer, J., Leventon, J., Newig, J., Schomerus, T., Vilsmaier, U., von Wehrden, H., Abernethy, P., Ives, C. D., Jager, N. W. & Lang, D. J. 2017. Leverage points for sustainability transformation. *Ambio,* 46**,** 30-39.

Achibet, M., Balev, S., Dutot, A. & Olivier, D. 2014. A Model of Road Network and Buildings Extension Co-evolution. *Procedia Computer Science,* 32**,** 828-833.

Adama, O. 2020. Slum upgrading in the era of World-Class city construction: the case of Lagos, Nigeria. *International Journal of Urban Sustainable Development,* 12**,** 219-235.

Adams, E. A. 2018a. Intra-urban inequalities in water access among households in Malawi's informal settlements: Toward pro-poor urban water policies in Africa. *Environmental Development,* 26**,** 34-42.

Adams, E. A. 2018b. Thirsty slums in African cities: household water insecurity in urban informal settlements of Lilongwe, Malawi. *International Journal of Water Resources Development,* 34**,** 869-887.

Adams, E. A., Byrns, S., Kumwenda, S., Quilliam, R., Mkandawire, T. & Price, H. 2022. Water journeys: Household water insecurity, health risks, and embodiment in slums and informal settlements. *Social Science & Medicine,* 313**,** 115394.

Adams, E. A., Sambu, D. & Smiley, S. L. 2019. Urban water supply in Sub-Saharan Africa: historical and emerging policies and institutional arrangements. *International Journal of Water Resources Development,* 35**,** 240-263.

Adams, E. A., Zulu, L. & Ouellette - Kray, Q. 2020. Community water governance for urban water security in the Global South: Status, lessons, and prospects. *Wiley Interdisciplinary Reviews: Water,* 7**,** e1466.

Adraoui, M., Diop, E. B., Ebnou Abdem, S. A., Azmi, R. & Chenal, J. 2024. Towards an Understanding of Hydraulic Sensitivity: Graph Theory Contributions to Water Distribution Analysis. *Water,* 16**,** 646.

Adusei, P. K., Oduro-Ofori, E., Amponsah, O. & Agyeman, K. O. 2018. Participatory incremental slum upgrading towards sustainability: an assessment of slum dwellers' willingness and ability to pay for utility services. *Environment, Development and Sustainability,* 20**,** 2501-2520.

Agathokleous, A., Christodoulou, C. & Christodoulou, S. E. 2017. Topological Robustness and Vulnerability Assessment of Water Distribution Networks. *Water Resources Management,* 31**,** 4007-4021.

Agreste, S., Meo, P. D., Fiumara, G., Piccione, G., Piccolo, S., Rosaci, D., Sarné, G. M. L. & Vasilakos, A. V. 2017. An Empirical Comparison of Algorithms to Find Communities in Directed Graphs and Their Application in Web Data Analytics. *IEEE Transactions on Big Data,* 3**,** 289-306.

Agyemang, F. S. K., Silva, E. & Poku-Boansi, M. 2019. Understanding the urban spatial structure of Sub-Saharan African cities using the case of urban development patterns of a Ghanaian city-region. *Habitat International,* 85**,** 21-33.

Ahimah-Agyakwah, S., Nketiah-Amponsah, E. & Agyire-Tettey, F. 2022. Urbanization and poverty in Sub-Saharan Africa: evidence from dynamic panel data analysis of selected urbanizing countries. *Cogent Economics & Finance,* 10**,** 2109282.

Ahlers, R., Schwartz, K. & Perez Guida, V. 2013. The myth of 'healthy' competition in the water sector: The case of small scale water providers. *Habitat International,* 38**,** 175-182.

Ahmad, N., Chester, M., Bondank, E., Arabi, M., Johnson, N. & Ruddell, B. L. 2022. A synthetic water distribution network model for urban resilience. *Sustainable and Resilient Infrastructure,* 7**,** 333-347.

Ahmed Saad, O., Anwar Fikry, M. & El-Sayed Hasan, A. 2019. Sustainable upgrading for informal areas. *Alexandria Engineering Journal,* 58**,** 237-249.

Ait-Kadi, M. 2016. Water for Development and Development for Water: Realizing the Sustainable Development Goals (SDGs) Vision. *Aquatic Procedia,* 6**,** 106-110.

Akbar, H. M. D., Minnery, J. R., van Horen, B. & Smith, P. 2007. Community water supply for the urban poor in developing countries: The case of Dhaka, Bangladesh. *Habitat International,* 31**,** 24-35.

Akbarzadeh, M., Memarmontazerin, S., Derrible, S. & Salehi Reihani, S. F. 2019. The role of travel demand and network centrality on the connectivity and resilience of an urban street system. *Transportation,* 46**,** 1127-1141.

Akbarzadeh, M., Memarmontazerin, S. & Soleimani, S. 2018. Where to look for power Laws in urban road networks? *Applied Network Science,* 3**,** 4.

Akelo, J. O. & Nzengya, D. M. 2023. How Household Water Use Characteristics and Sense of Water Kiosk Ownership Influence Financial Sustainability of Community-Managed Borehole Water Kiosk Service in Kisumu County, Kenya. *African Journal of Empirical Research,* 4**,** 630-645.

Akinyemi, F. O. & Bigirimana, F. 2012. A Spatial Analysis of Poverty in Kigali, Rwanda using indicators of household living standard. *Rwanda Journal,* 26**,** 3-22.

Alaazi, D. A. & Aganah, G. A. M. 2020. Understanding the slum–health conundrum in sub-Saharan Africa: a proposal for a rights-based approach to health promotion in slums. *Global Health Promotion,* 27**,** 65-72.

Alba, R. & Bruns, A. 2022. First-Class but not for Long: Heterogeneous Infrastructure and Water Bricolage in Accra's Kiosk Compounds. *Urban Forum,* 33**,** 129-151.

Alizadeh, T. & Prasad, D. 2024. The right to the smart city in the Global South: A research agenda. *Urban Studies,* 61**,** 426-444.

Allen, A., Dávila, J. D. & Hofmann, P. 2006. *Governance of water and sanitation services for the peri-urban poor: A framework for understanding and action in metropolitan regions*, The Development Planning Unit, University College London.

Almeile, A. M., Chipulu, M., Ojiako, U., Vahidi, R. & Marshall, A. 2024. Project-focussed literature on public-private partnership (PPP) in developing countries: a critical review. *Production Planning & Control,* 35**,** 683-710.

Almheiri, Z., Meguid, M. & Zayed, T. 2020. Intelligent Approaches for Predicting Failure of Water Mains. *Journal of Pipeline Systems Engineering and Practice,* 11**,** 04020044.

Alobaydi, D. & Rashid, M. 2024. The evolution of street structures: A morphological study. *Journal of Engineering,* 30**,** 203-220.

Alvarado, N. A. 2022. Migrant Politics in the Urban Global South: The Political Work of Nicaraguan Migrants to Acquire Urban Rights in Costa Rica. *Geopolitics,* 27**,** 1180-1204.

Amegah, A. K. 2021. Slum decay in Sub-Saharan Africa: Context, environmental pollution challenges, and impact on dweller's health. *Environ Epidemiol,* 5**,** e158.

American Red Cross 2019. Administrative Wards in Kenya 1450. *In:* CROSS, A. R. (ed.).

Ameyaw, E. E. & Chan, A. P. 2015. Evaluating key risk factors for PPP water projects in Ghana: a Delphi study. *Journal of Facilities Management,* 13**,** 133-155.

Ameyaw, E. E. & Chan, A. P. 2016. Critical success factors for public-private partnership in water supply projects. *Facilities,* 34**,** 124-160.

Ananga, E. O., Njoh, A. J., Adams, E. A. & Peaden, C. 2019. Factors impeding efficient functionality of community operated urban water schemes: the experience of Kisumu, Kenya. *Sustainable Water Resources Management,* 5**,** 1143-1153.

Andavarapu, D. & Edelman, D. J. 2013. Evolution of slum redevelopment policy. *Current urban studies,* 1**,** 185.

Andersson, H. 2017. *Colonial Urban Legacies: An analysis of socio-spatial structures in Accra, Ghana.* Bachelor, Uppsala University.

Andres, L. A., Guasch, J. L., Haven, T. & Foster, V. 2008. *The impact of private sector participation in infrastructure: lights, shadows, and the road ahead*, World Bank Publications.

Añez, J., De La Barra, T. & Pérez, B. 1996. Dual graph representation of transport networks. *Transportation Research Part B: Methodological,* 30**,** 209-216.

Annamalai, T., Devkar, G., Mahalingam, A., Benjamin, S., Rajan, S. & Deep, A. 2016. What is the evidence on top-down and bottom-up approaches in improving access to water, sanitation and electricity services in low-income or informal settlements. *Social Science Research Unit, UCL Institute of Education, University College London***,** 122.

Anselin, L. 1995. Local indicators of spatial association—LISA. *Geographical analysis,* 27**,** 93-115.

Antos, S. E., Lall, S. V. & Lozano-Gracia, N. 2016. The morphology of African cities. *World Bank Policy Research Working Paper*.

Antunes, M. & Martins, R. 2020. Determinants of access to improved water sources: Meeting the MDGs. *Utilities Policy,* 63**,** 101019.

Apparicio, P., Gelb, J., Dubé, A.-S., Kingham, S., Gauvin, L. & Robitaille, É. 2017. The approaches to measuring the potential spatial access to urban health services revisited: distance types and aggregation-error issues. *International Journal of Health Geographics,* 16**,** 32.

Araújo de Oliveira, V. M. 2022a. The Elements of Urban Form. *In:* ARAúJO DE OLIVEIRA, V. M. (ed.) *Urban Morphology: An Introduction to the Study of the Physical Form of Cities*. Cham: Springer International Publishing.

Araújo de Oliveira, V. M. 2022b. The Study of Urban Form: Different Approaches. *In:* ARAúJO DE OLIVEIRA, V. M. (ed.) *Urban Morphology: An Introduction to the Study of the Physical Form of Cities*. Cham: Springer International Publishing.

Armah, F. A., Ekumah, B., Yawson, D. O., Odoi, J. O., Afitiri, A.-R. & Nyieku, F. E. 2018. Access to improved water and sanitation in sub-Saharan Africa in a quarter century. *Heliyon,* 4**,** e00931.

Artico, I., Smolyarenko, I., Vinciotti, V. & Wit, E. C. 2020. How rare are power-law networks really? *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 476**,** 20190742.

Aşchilean, I., Iliescu, M., Ciont, N. & Giurca, I. 2018. The Unfavourable Impact of Street Traffic on Water Distribution Pipelines. *Water,* 10**,** 1086.

Asumadu, G., Quaigrain, R., Owusu-Manu, D., Edwards, D. J., Oduro-Ofori, E. & Dapaah, S. M. 2023. Analysis of urban slum infrastructure projects financing in Ghana: A closer look at traditional and innovative financing mechanisms. *World Development Perspectives,* 30**,** 100505.

Auerbach, A. M., LeBas, A., Post, A. E. & Weitz-Shapiro, R. 2018. State, Society, and Informality in Cities of the Global South. *Studies in Comparative International Development,* 53**,** 261-280.

Azunre, G. A., Amponsah, O., Takyi, S. A. & Mensah, H. 2021. Informality-sustainable city nexus: The place of informality in advancing sustainable Ghanaian cities. *Sustainable Cities and Society,* 67**,** 102707.

Azunre, G. A., Amponsah, O., Takyi, S. A., Mensah, H. & Braimah, I. 2022. Urban informalities in sub-Saharan Africa (SSA): A solution for or barrier against sustainable city development. *World Development,* 152**,** 105782.

Badhrudeen, M., Derrible, S., Verma, T., Kermanshah, A. & Furno, A. 2022. A Geometric Classification of World Urban Road Networks. *Urban Science,* 6**,** 11.

Baffoe, G., Ahmad, S. & Bhandari, R. 2020a. The road to sustainable Kigali: A contextualized analysis of the challenges. *Cities,* 105**,** 102838.

Baffoe, G., Malonza, J., Manirakiza, V. & Mugabe, L. 2020b. Understanding the Concept of Neighbourhood in Kigali City, Rwanda. *Sustainability,* 12**,** 1555.

Bain, R. E., Gundry, S. W., Wright, J. A., Yang, H., Pedley, S. & Bartram, J. K. 2012. Accounting for water quality in monitoring access to safe drinking-water as part of the Millennium Development Goals: lessons from five countries. *Bulletin of the World Health Organization,* 90**,** 228-235.

Banda, E. M. & Mwale, F. D. 2018. Utility performance in supplying water to informal settlements: A case study from Malawi. *Utilities Policy,* 55**,** 151-157.

Banks, N., Lombard, M. & Mitlin, D. 2020. Urban Informality as a Site of Critical Analysis. *The Journal of Development Studies,* 56**,** 223-238.

Barabási, A.-L., Gulbahce, N. & Loscalzo, J. 2011. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics,* 12**,** 56-68.

Barrat, A., Barthélemy, M., Pastor-Satorras, R. & Vespignani, A. 2004. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences,* 101**,** 3747-3752.

Barthélemy, M. 2011. Spatial networks. *Physics Reports,* 499**,** 1-101.

Bartram, J., Brocklehurst, C., Fisher, M. B., Luyendijk, R., Hossain, R., Wardlaw, T. & Gordon, B. 2014. Global monitoring of water supply and sanitation: history, methods and future challenges. *International journal of environmental research and public health,* 11**,** 8137-8165.

Bastia, T. 2015. Transnational migration and urban informality: Ethnicity in Buenos Aires' informal settlements. *Urban Studies,* 52**,** 1810-1825.

Batty, M. 2012. Building a science of cities. *Cities,* 29**,** S9-S16.

Bedi, P. & Sharma, C. 2016. Community detection in social networks. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 6**,** 115-135.

Benken, E. E. 2017. *Nowhere to go: informal settlement eradication in Kigali, Rwanda.* University of Louisville.

Bentes, I., Afonso, L., Varum, H., Pinto, J., Varajão, J., Duarte, A. & Agarwal, J. 2011. A new tool to assess water pipe networks vulnerability and robustness. *Engineering Failure Analysis,* 18**,** 1637-1644.

Bigon, L. & Njoh, A. J. 2015. The Toponymic Inscription Problematic in Urban Sub-Saharan Africa: From Colonial to Postcolonial Times. *Journal of Asian and African Studies,* 50**,** 25-40.

Bittencourt, T. A. & Giannotti, M. 2023. Evaluating the accessibility and availability of public services to reduce inequalities in everyday mobility. *Transportation Research Part A: Policy and Practice,* 177**,** 103833.

Blagus, N., Šubelj, L. & Bajec, M. 2014. Assessing the effectiveness of real-world network simplification. *Physica A: Statistical Mechanics and its Applications,* 413**,** 134-146.

Blanc, D. L., Freire, C. & Vierros, M. 2017. Mapping the Linkages between Oceans and Other Sustainable Development Goals. *United Nations*.

Bland, J. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*.

Blöcker, C., Nieves, J. C. & Rosvall, M. 2022. Map equation centrality: community-aware centrality based on the map equation. *Applied Network Science,* 7**,** 56.

Blondel, V. D., Gajardo, A., Heymans, M., Senellart, P. & Dooren, P. V. 2004. A Measure of Similarity between Graph Vertices: Applications to Synonym Extraction and Web Searching. *SIAM Review,* 46**,** 647-666.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment,* 2008**,** P10008.

Boakye-Ansah, A. S., Schwartz, K. & Zwarteveen, M. 2019. From Rowdy Cartels to Organized Ones? The Transfer of Power in Urban Water Supply in Kenya. *The European Journal of Development Research,* 31**,** 1246-1262.

Boakye-Ansah, A. S., Schwartz, K. & Zwarteveen, M. 2022. Service differentiation as an improvement strategy for access to water in urban low-income areas: evidence from three Kenyan cities. *International Journal of Water Resources Development,* 38**,** 766-782.

Boccaletti, S., Bianconi, G., Criado, R., del Genio, C. I., Gómez-Gardeñes, J., Romance, M., Sendiña-Nadal, I., Wang, Z. & Zanin, M. 2014. The structure and dynamics of multilayer networks. *Physics Reports,* 544**,** 1-122.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D. U. 2006. Complex networks: Structure and dynamics. *Physics Reports,* 424**,** 175-308.

Boeing, G. 2017. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems,* 65**,** 126-139.

Boeing, G. 2018. Planarity and street network representation in urban form analysis. *Environment and Planning B: Urban Analytics and City Science,* 47**,** 855-869.

Bolin, A. 2019. A Ghost Map of Kigali. *Anthropology and Humanism,* 44**,** 144-155.

Bolton, L. 2020. 'Sites and services', and in-situ slum upgrading. *K4D Helpdesk Report.* Brighton, UK.

Bondarenko, M., Kerr, D., Sorichetta, A. & Tatem, A. 2020. Census/projection-disaggregated gridded population datasets, adjusted to match the corresponding UNPD 2020 estimates, for 51 countries across sub-Saharan Africa using building footprints.

Boschmann, E. E. & Kwan, M.-P. 2008. Toward Socially Sustainable Urban Transportation: Progress and Potentials. *International Journal of Sustainable Transportation,* 2**,** 138-157.

Bramson, A. 2022. Neighborhood discovery via augmented network community structure. *Applied Network Science,* 7**,** 51.

Brelsford, C. & Martin, T. 2021. The Mathematical Foundations of the Science of Cities. *In:* SRIRAMAN, B. (ed.) *Handbook of the Mathematics of the Arts and Sciences.* Cham: Springer International Publishing.

Brelsford, C., Martin, T., Hand, J. & Bettencourt, L. M. A. 2018. Toward cities without slums: Topology and the spatial evolution of neighborhoods. *Science Advances,* 4**,** eaar4644.

Brenner, N. 2009. What is critical urban theory? *City,* 13**,** 198-207.

Brenner, N. & Schmid, C. 2015. Towards a new epistemology of the urban? *City,* 19**,** 151-182.

Brentan, B., Campbell, E., Goulart, T., Manzi, D., Meirelles, G., Herrera, M., Izquierdo, J. & Luvizotto, E. 2018. Social Network Community Detection and Hybrid Optimization for Dividing Water Supply into District Metered Areas. *Journal of Water Resources Planning and Management,* 144**,** 04018020.

Brentan, B. M., Campbell, E., Meirelles, G. L., Luvizotto, E. & Izquierdo, J. 2017. Social Network Community Detection for DMA Creation: Criteria Analysis through Multilevel Optimization. *Mathematical Problems in Engineering,* 2017**,** 9053238.

Breuer, J. H. P., Friesen, J., Taubenböck, H., Wurm, M. & Pelz, P. F. 2024. The unseen population: Do we underestimate slum dwellers in cities of the Global South? *Habitat International,* 148**,** 103056.

Broido, A. D. & Clauset, A. 2019. Scale-free networks are rare. *Nature Communications,* 10**,** 1017.

Brown-Luthango, M., Reyes, E. & Gubevu, M. 2017. Informal settlement upgrading and safety: experiences from Cape Town, South Africa. *Journal of Housing and the Built Environment,* 32**,** 471-493.

Bryant Jr, J. & Delamater, P. L. 2019. Examination of spatial accessibility at micro- and macro-levels using the enhanced two-step floating catchment area (E2SFCA) method. *Annals of GIS,* 25**,** 219-229.

Bugden, D. & Stedman, R. 2019. Place and behavior: The role of accessibility. *Journal of Environmental Psychology,* 63**,** 109-117.

Buhl, C., Gautrais, J., Reeves, N., Solé, R. V., Valverde, S., Kuntz, P. & Theraulaz, G. 2006. Topological patterns in street networks of self-organizedurban settlements. *The European Physical Journal B - Condensed Matter and ComplexSystems,* 49**,** 513-522.

Bunke, H. & Allermann, G. 1983. Inexact graph matching for structural pattern recognition. *Pattern Recognition Letters,* 1**,** 245-253.

Cain, A. 2018. Informal water markets and community management in peri-urban Luanda, Angola. *Water International,* 43**,** 205-216.

Caldeira, T. P. 2017. Peripheral urbanization: Autoconstruction, transversal logics, and politics in cities of the global south. *Environment and Planning D: Society and Space,* 35**,** 3-20.

Campbell, E., Ayala-Cabrera, D., Izquierdo, J., Pérez-García, R. & Tavera, M. 2014. Water Supply Network Sectorization Based on Social Networks Community Detection Algorithms. *Procedia Engineering,* 89**,** 1208-1215.

Cardillo, A., Scellato, S., Latora, V. & Porta, S. 2006. Structural properties of planar graphs of urban street patterns. *Physical Review E,* 73**,** 066107.

Casali, Y. & Heinimann, H. R. 2019. A topological analysis of growth in the Zurich road network. *Computers, Environment and Urban Systems,* 75**,** 244-253.

Cassivi, A., Guilherme, S., Bain, R., Tilley, E., Waygood, E. O. D. & Dorea, C. 2019. Drinking water accessibility and quantity in low and middle-income countries: A systematic review. *International Journal of Hygiene and Environmental Health,* 222**,** 1011-1020.

Cassivi, A., Johnston, R., Waygood, E. O. D. & Dorea, C. C. 2018. Access to drinking water: time matters. *Journal of Water and Health,* 16**,** 661-666.

Cassivi, A., Tilley, E., Waygood, E. O. D. & Dorea, C. 2021. Evaluating self-reported measures and alternatives to monitor access to drinking water: A case study in Malawi. *Science of The Total Environment,* 750**,** 141516.

Castells-Quintana, D. & Wenban-Smith, H. 2020. Population Dynamics, Urbanisation without Growth, and the Rise of Megacities. *The Journal of Development Studies,* 56**,** 1663-1682.

Castro, V. & Morel, A. 2008. Can delegated management help water utilities improve services to informal settlements? *Waterlines***,** 289-306.

Center for International Earth Science Information Network - CIESIN - Columbia University 2018. Gridded Population of the World, Version 4 (GPWv4): Population Density Adjusted to Match 2015 Revision UN WPP Country Totals, Revision 11. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC).

Cesar, S. 2019. Privatization of water: evaluating its performance in the developing world. *Annals of Public and Cooperative Economics,* 90**,** 5-23.

CGK 2022. Kisumu County Integrated Development Plan (2023-2027).

Chakraborty, S. 2022. Periurban Water: Recognizing the Margins for Sustainable Urban Water Futures. *In:* LEAL FILHO, W., AZUL, A. M., BRANDLI, L., LANGE SALVIA, A. & WALL, T. (eds.) *Clean Water and Sanitation.* Cham: Springer International Publishing.

Charmes, J. 2012. The informal economy worldwide: Trends and characteristics. *Margin: the journal of applied economic research,* 6**,** 103-132.

Chattopadhyay, S., Chakraborty, T., Ghosh, K. & Das, A. K. 2021. Modified Lomax model: a heavy-tailed distribution for fitting large-scale real-world complex networks. *Social Network Analysis and Mining,* 11**,** 43.

Chegini, T. & Li, H. Y. 2022. An algorithm for deriving the topology of belowground urban stormwater networks. *Hydrol. Earth Syst. Sci.,* 26**,** 4279-4300.

Chen, B. Y., Yuan, H., Li, Q., Wang, D., Shaw, S.-L., Chen, H.-P. & Lam, W. H. K. 2017. Measuring place-based accessibility under travel time uncertainty. *International Journal of Geographical Information Science,* 31**,** 783-804.

Chen, J., Saad, Y. & Zhang, Z. 2022a. Graph coarsening: from scientific computing to machine learning. *SeMA Journal,* 79**,** 187-223.

Chen, M., Kuzmin, K. & Szymanski, B. K. 2014. Community detection via maximization of modularity and its variants. *IEEE Transactions on Computational Social Systems,* 1**,** 46-65.

Chen, M. A. 2012. The informal economy: Definitions, theories and policies. *WIEGO Working Paper* Manchester, United Kingdom: Women in Informal Employment: Globalizing and Organizing (WIEGO).

Chen, X. 2019. Enhancing the Two-Step Floating Catchment Area Model for Community Food Access Mapping: Case of the Supplemental Nutrition Assistance Program. *The Professional Geographer,* 71**,** 668-680.

Chen, X. & Jia, P. 2019. A comparative analysis of accessibility measures by the two-step floating catchment area (2SFCA) method. *International Journal of Geographical Information Science,* 33**,** 1739-1758.

Chen, Y., Liu, Z. & Zhou, B.-B. 2022b. Population-environment dynamics across world's top 100 urban agglomerations: With implications for transitioning toward global urban sustainability. *Journal of Environmental Management,* 319**,** 115630.

Cheng, X. & Scherpen, J. M. A. 2021. Model Reduction Methods for Complex Network Systems. *Annual Review of Control, Robotics, and Autonomous Systems,* 4**,** 425-453.

Chenoweth, J. & Bird, J. 2018. *The business of water and sustainable development*, Routledge.

Chidambaram, S. 2020. How do institutions and infrastructure affect mobilization around public toilets vs. piped water? Examining intra-slum patterns of collective action in Delhi, India. *World Development,* 132**,** 104984.

Chintalapudi, S. R. & Prasad, M. K. A survey on community detection algorithms in large scale real world networks. 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015. IEEE, 1323-1327.

Chitonge, H. 2014. Cities Beyond Networks: The Status of Water Services for the Urban Poor in African Cities. *African Studies,* 73**,** 58-83.

Chittor Sundaram, R., Naghizade, E., Borovica-Gajic, R. & Tomko, M. 2020. Harnessing spatio-temporal patterns in data for nominal attribute imputation. *Transactions in GIS,* 24**,** 1001-1032.

Cities Alliance 2021a. Community-led mechanisms: For upgrading informal settlements and informal work in southern cities. *Global Review Series on Informality.*

Cities Alliance 2021b. COVID-19 and informality: Good practices for reducing risk and enhancing resilience. *Global Review Series on Informality.*

Cities Alliance 2021c. An international review of slum upgrading practices: Lessons learned and remaining challenges of projects and programs. *Global Review Series on Informality.*

Cities Alliance. n.d. *Slums and Slum Upgrading* [Online]. Available: https://www.citiesalliance.org/themes/slums-and-slum-upgrading [Accessed].

City of Kigali 2020a. Kigali Master Plan 2050: Analysis and Vision. Kigali, Rwanda: Surbana Jurong Consultants Pte Ltd.

City of Kigali 2020b. Kigali Master Plan 2050: Main Report. Kigali, Rwanda.

City of Kigali 2020c. Kigali Master Plan 2050: Transport Plan. Kigali, Rwanda.

City of Kigali 2020d. Kigali Master Plan 2050: Zoning Regulations. Kigali, Rwanda.

Clark, J. & Holton, D. A. 1991. *A first look at graph theory*, World Scientific.

Cobbinah, P. B. & Niminga-Beka, R. 2017. Urbanisation in Ghana: Residential land use under siege in Kumasi central. *Cities,* 60**,** 388-401.

Comber, A. J., Brunsdon, C. F. & Farmer, C. J. Q. 2012. Community detection in spatial networks: Inferring land use from a planar graph of land cover objects. *International Journal of Applied Earth Observation and Geoinformation,* 18**,** 274-282.

Contzen, N. & Marks, S. J. 2018. Increasing the regular use of safe water kiosk through collective psychological ownership: A mediation analysis. *Journal of Environmental Psychology,* 57**,** 45-52.

Coopman, A., Osborn, D., Ullah, F., Auckland, E. & Long, G. Seeing the whole: implementing the SDGs in an integrated and coherent way. Stakeholder Forum. London. UK, 2016.

Coşkun, M. & Koyutürk, M. 2021. Node similarity-based graph convolution for link prediction in biological networks. *Bioinformatics,* 37**,** 4501-4508.

County of Kisumu 2013. The Integrated Strategic Urban Development Plan: Planning for Kisumu's Future Kisumu, Kenya.

County of Kisumu 2020. Kisumu Sustainable Mobility Plan. Kisumu, Kenya.

Coupette, C. & Vreeken, J. 2021. Graph Similarity Description: How Are These Graphs Similar? *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining.* Virtual Event, Singapore: Association for Computing Machinery.

Coutard, O. 2008. Placing splintering urbanism: Introduction. *Geoforum,* 39**,** 1815-1820.

Coutard, O. & Rutherford, J. 2015. *Beyond the networked city: Infrastructure reconfigurations and urban change in the North and South*, Routledge.

Criqui, L. 2018. L'espace urbain : espace dual ou mosaïque ? Inégalités et fragmentation urbaine dans les villes des Sud. *Revue internationale et stratégique,* 112**,** 129-138.

Croese, S., Dominique, M. & Raimundo, I. M. 2021. Co-producing urban knowledge in Angola and Mozambique: towards meeting SDG 11. *npj Urban Sustainability,* 1**,** 8.

Cronin, V. & Guthrie, P. 2011. Alternative approaches to slum upgrading in Kibera, Nairobi. *Proceedings of the Institution of Civil Engineers - Urban Design and Planning,* 164**,** 129-139.

Crow, B., Davies, J., Paterson, S. & Miles, J. 2013. Using GPS and recall to understand water collection in Kenyan informal settlements. *Water International,* 38**,** 43-60.

Csardi, G. & Nepusz, T. 2006. The igraph software package for complex network research. *InterJournal, complex systems,* 1695**,** 1-9.

D'Ambrosio, C., Lodi, A., Wiese, S. & Bragalli, C. 2015. Mathematical programming techniques in water network optimization. *European Journal of Operational Research,* 243**,** 774-788.

Dai, D. 2010. Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. *Health & Place,* 16**,** 1038-1052.

Dakyaga, F., Ahmed, A. & Sillim, M. L. 2021. Governing Ourselves for Sustainability: Everyday Ingenuities in the Governance of Water Infrastructure in the Informal Settlements of Dar es Salaam. *Urban Forum,* 32**,** 111-129.

Dalvi, M. Q. & Martin, K. M. 1976. The measurement of accessibility: Some preliminary results. *Transportation,* 5**,** 17-42.

Daly, S. W., Lowe, J., Hornsby, G. M. & Harris, A. R. 2021. Multiple water source use in low- and middle-income countries: a systematic review. *Journal of Water and Health,* 19**,** 370-392.

Daniel, C. B., Mathew, S. & Subbarayan, S. 2021. GIS-Based Study on the Association Between Road Centrality and Socio-demographic Parameters: a Case Study. *Journal of Geovisualization and Spatial Analysis,* 6**,** 1.

Danon, L., Díaz-Guilera, A., Duch, J. & Arenas, A. 2005. Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment,* 2005**,** P09008.

Dao, V. L., Bothorel, C. & Lenca, P. 2020. Community structure: A comparative evaluation of community detection methods. *Network Science,* 8**,** 1-41.

Datta, D. 2024. blandr: Bland-Altman Method Comparison for R, version 0.6.0.

Davis, M. 2004. Planet of slums. *New left review,* 26**,** 5.

de Castro Mazarro, A. 2023. Situating Slums in Hegemonic Urban Discourse: A Historiography of English-Language Architecture and Planning Journals. *Journal of Urban History,* 49**,** 533-551.

Debón, A., Carrión, A., Cabrera, E. & Solano, H. 2010. Comparing risk of failure models in water supply networks using ROC curves. *Reliability Engineering & System Safety,* 95**,** 43-48.

Delaunois, F., Tosar, F. & Vitry, V. 2014. Corrosion behaviour and biocorrosion of galvanized steel water distribution systems. *Bioelectrochemistry,* 97**,** 110-119.

Delignette-Muller, M. L. & Dutang, C. 2015. fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software,* 64**,** 1 - 34.

Derudder, B. & Neal, Z. 2018. Uncovering Links Between Urban Studies and Network Science. *Networks and Spatial Economics,* 18**,** 441-446.

Deshpande, A., Miller-Petrie, M. K., Lindstedt, P. A., Baumann, M. M., Johnson, K. B., Blacker, B. F., Abbastabar, H., Abd-Allah, F., Abdelalim, A. & Abdollahpour, I. 2020. Mapping geographical inequalities in access to drinking water and sanitation facilities in low-income and middle-income countries, 2000–17. *The Lancet Global Health,* 8**,** e1162-e1185.

Despalatović, L., Vojković, T. & Vukicević, D. Community structure in networks: Girvan-Newman algorithm improvement. 2014 37th international convention on information and communication technology, electronics and microelectronics (MIPRO), 2014. IEEE, 997-1002.

Deuerlein, J. W. 2008. Decomposition Model of a General Water Supply Network Graph. *Journal of Hydraulic Engineering,* 134**,** 822-832.

Devi, A. & Bostoen, K. 2009. Extending the critical aspects of the water access indicator using East Africa as an example. *International Journal of Environmental Health Research,* 19**,** 329-341.

Di Virgilio, M. M. 2023. Urbicide: Towards a Conceptualization. *In:* CARRIóN MENA, F. & CEPEDA PICO, P. (eds.) *Urbicide: The Death of the City.* Cham: Springer International Publishing.

Dias, M. D., Petronetto, F., Valdivia, P. & Nonato, L. G. Graph Spectral Filtering for Network Simplification. 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 29 Oct.-1 Nov. 2018 2018. 345-352.

Dingil, A. E., Schweizer, J., Rupi, F. & Stasiskiene, Z. 2018. Transport indicator analysis and comparison of 151 urban areas, based on open source data. *European Transport Research Review,* 10**,** 58.

Dixit, M. & Sivakumar, A. 2020. Capturing the impact of individual characteristics on transport accessibility and equity analysis. *Transportation Research Part D: Transport and Environment,* 87**,** 102473.

Dong, X., Thanou, D., Toni, L., Bronstein, M. & Frossard, P. 2020. Graph signal processing for machine learning: A review and new perspectives. *IEEE Signal processing magazine,* 37**,** 117-127.

Dongzagla, A., Dordaa, F. & Agbenyo, F. 2022. Spatial inequality in safely managed water access in Ghana. *Journal of Water, Sanitation and Hygiene for Development,* 12**,** 869-882.

Dongzagla, A., Nunbogu, A. M. & Fielmua, N. 2020. Does self-reported water collection time differ from observed water collection time? Evidence from the Upper West Region of Ghana. *Journal of Water, Sanitation and Hygiene for Development,* 10**,** 357-365.

Dos Santos, S., Adams, E. A., Neville, G., Wada, Y., de Sherbinin, A., Mullin Bernhardt, E. & Adamo, S. B. 2017. Urban growth and water access in sub-Saharan Africa: Progress, challenges, and emerging research directions. *Science of The Total Environment,* 607-608**,** 497-508.

Dovey, K., Shafique, T., Oostrum, M. v. & Chatterjee, I. 2021a. Informal settlement is not a euphemism for 'slum': what's at stake beyond the language? *International Development Planning Review,* 43**,** 139-150.

Dovey, K., Shafique, T., van Oostrum, M. & Chatterjee, I. 2021b. Informal settlement is not a euphemism for ‘slum’: what’s at stake beyond the language? *International Development Planning Review,* 43**,** 139-150.

Dovey, K., van Oostrum, M., Chatterjee, I. & Shafique, T. 2020. Towards a morphogenesis of informal settlements. *Habitat International,* 104**,** 102240.

Dufitimana, E. & Niyonzima, T. 2023. Leveraging the Potential of Convolutional Neural Network and Satellite Images to Map Informal Settlements in Urban Settings of the City of Kigali, Rwanda. *Rwanda Journal of Engineering, Science, Technology and Environment,* 5.

Dunn, S., Fu, G., Wilkinson, S. & Dawson, R. 2013. Network theory for infrastructure systems modelling. *Proceedings of the Institution of Civil Engineers - Engineering Sustainability,* 166**,** 281-292.

Dupuy, G. 2011. Fracture et dépendance: l'enfer des réseaux? *Flux,* 83**,** 6-23.

Edith, E. M., Maurice, O. & Okuku, M. T. 2019. In-Situ Slum Upgrading Programs on Improved Livelihoods in Manyatta Informal Settlements, Kisumu City, Kenya. *International Journal of Social and Development Concerns,* 7**,** 38-54.

Erdős, P. & Rényi, A. 1960. On the evolution of random graphs. *Publ. math. inst. hung. acad. sci,* 5**,** 17-60.

Ersoy, O., Hurter, C., Paulovich, F., Cantareiro, G. & Telea, A. 2011. Skeleton-based edge bundling for graph visualization. *IEEE transactions on visualization and computer graphics,* 17**,** 2364-2373.

Esmail, S. & Corburn, J. 2020. Struggles to remain in Kigali’s “unplanned” settlements: the case of Bannyahe. *Environment and Urbanization,* 32**,** 19-36.

Ezeh, A., Oyebode, O., Satterthwaite, D., Chen, Y. F., Ndugwa, R., Sartori, J., Mberu, B., Melendez-Torres, G. J., Haregu, T., Watson, S. I., Caiaffa, W., Capon, A. & Lilford, R. J. 2017. The history, geography, and sociology of slums and the health problems of people who live in slums. *Lancet,* 389**,** 547-558.

Facebook Connectivity Lab and CIESIN-Columbia University 2016. High Resolution Settlement Layer.

Falcone, M., Salvinelli, C., Bah, M. & Thomas, E. 2023. Effectiveness of a water-vending kiosk intervention toward household water quality and surveyed water security in Freetown, Sierra Leone. *Science of The Total Environment,* 875**,** 162447.

FAO 2020. Land cover classification (Africa and Near East - Annual - 100m) - WaPOR v2.

Farage, C., Edler, D., Eklöf, A., Rosvall, M. & Pilosof, S. 2021. Identifying flow modules in ecological networks using Infomap. *Methods in Ecology and Evolution,* 12**,** 778-786.

Farmani, R. & Butler, D. 2014. Implications of Urban Form on Water Distribution Systems Performance. *Water Resources Management,* 28**,** 83-97.

Feeny, S. 2020. Transitioning from the MDGs to the SDGs: Lessons learnt? *Moving from the Millennium to the Sustainable Development Goals: Lessons and recommendations***,** 343-351.

Fix, M. & Arantes, P. F. 2022. On urban studies in Brazil: The favela, uneven urbanisation and beyond. *Urban Studies,* 59**,** 893-916.

Flores Fernandez, R. A. 2011. Physical and Spatial Characteristics of Slum Territories—Vulnerable to Natural Disasters. *Les Cahiers d'Afrique de l'Est/The East African Review*, 5-22.

Fornito, A., Zalesky, A. & Bullmore, E. 2016a. Chapter 4 - Node Degree and Strength. *Fundamentals of Brain Network Analysis.* San Diego: Academic Press.

Fornito, A., Zalesky, A. & Bullmore, E. 2016b. *Fundamentals of brain network analysis*, Academic press.

Fortunato, S. 2010. Community detection in graphs. *Physics Reports,* 486, 75-174.

Fortunato, S. & Barthélemy, M. 2007. Resolution limit in community detection. *Proceedings of the National Academy of Sciences,* 104, 36-41.

Fortunato, S. & Hric, D. 2016. Community detection in networks: A user guide. *Physics reports,* 659, 1-44.

Fortunato, S. & Newman, M. E. J. 2022. 20 years of network community detection. *Nature Physics,* 18, 848-850.

Foster, T. 2013. Predictors of Sustainability for Community-Managed Handpumps in Sub-Saharan Africa: Evidence from Liberia, Sierra Leone, and Uganda. *Environmental Science & Technology,* 47, 12037-12046.

Franco, I. B., Chatterji, T., Derbyshire, E. & Tracey, J. 2019. *Actioning the global goals for local impact*, Springer.

Fransen, K., Neutens, T., De Maeyer, P. & Deruyter, G. 2015. A commuter-based two-step floating catchment area method for measuring spatial accessibility of daycare centers. *Health & Place,* 32, 65-73.

Freeman, L. C. 1977. A set of measures of centrality based on betweenness. *Sociometry*, 35-41.

Freire, S., MacManus, K., Pesaresi, M., Doxsey-Whitfield, E. & Mills, J. 2016. Development of new open and free multi-temporal global population grids at 250 m resolution. *Population,* 250, 33.

Gadiaga, A., Abbott, T., Chamberlain, H., Lazar, A., Darin, E. & Tatem, A. 2023. Gridded disaggregated population estimates for Kenya, version 2.0.

Gandy, M. 2004. Rethinking urban metabolism: water, space and the modern city. *City,* 8, 363-379.

Gates, K. M., Henry, T., Steinley, D. & Fair, D. A. 2016. A Monte Carlo Evaluation of Weighted Community Detection Algorithms. *Frontiers in Neuroinformatics,* 10.

Getis, A. & Ord, J. K. 1992. The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis,* 24, 189-206.

Geurs, K. T. & van Wee, B. 2004. Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography,* 12, 127-140.

Gharaee, Z., Kowshik, S., Stromann, O. & Felsberg, M. 2021. Graph representation learning for road type classification. *Pattern Recognition,* 120, 108174.

Ghasemian, A., Hosseinmardi, H. & Clauset, A. 2019. Evaluating overfit and underfit in models of network community structure. *IEEE Transactions on Knowledge and Data Engineering,* 32, 1722-1735.

Gheisi, A., Forsyth, M. & Naser, G. 2016. Water distribution systems reliability: A review of research literature. *Journal of Water Resources Planning and Management,* 142**,** 04016047.

Ghosn, M., Dueñas-Osorio, L., Frangopol, D., McAllister, T., Bocchini, P., Manuel, L., Ellingwood, B., Arangio, S., Bontempi, F. & Shah, M. 2016. Performance indicators for structural systems and infrastructure networks. *Journal of Structural Engineering,* 142**,** F4016003.

Gilbert, A. 2007. The Return of the Slum: Does Language Matter? *International Journal of Urban and Regional Research,* 31**,** 697-713.

Gilson, G. G., Garrick, D., Cersey Ochieng, O., Ochieng Okoth, E. & Morara Ogendi, G. 2025. Can Informal Water Vendors be Trusted? The Evolution of Informality in Kisumu's Delegated Management Model. *The Journal of Development Studies***,** 1-18.

Girvan, M. & Newman, M. E. J. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences,* 99**,** 7821-7826.

Giudicianni, C., Di Nardo, A., Di Natale, M., Greco, R., Santonastaso, G. F. & Scala, A. 2018. Topological Taxonomy of Water Distribution Networks. *Water,* 10**,** 444.

Giustolisi, O., Ridolfi, L. & Simone, A. 2019. Tailoring Centrality Metrics for Water Distribution Networks. *Water Resources Research,* 55**,** 2348-2369.

Godfrey, S. & Howard, G. 2005a. *Water Safety Plans - Book 1: Planning water safety management for urban piped water supplies in developing countries*, WEDC, Loughborough University.

Godfrey, S. & Howard, G. 2005b. *Water Safety Plans - Book 2: Planning water safety management for urban piped water supplies in developing countries*, WEDC, Loughborough University.

GoK 2002. The Water Act 2002. *In:* PRINTER., N. G. (ed.).

GoK 2004. Kenya slum upgrading Strategy: Implementation strategy 2005-2020. Kiusmu, Kenya.

Gollin, D., Jedwab, R. & Vollrath, D. 2016. Urbanization with and without industrialization. *Journal of Economic Growth,* 21**,** 35-70.

GoR 2012. Rwanda Vision 2020. Kigali, Rwanda.

Gori, M., Monfardini, G. & Scarselli, F. A new model for learning in graph domains. Proceedings. 2005 IEEE international joint conference on neural networks, 2005., 2005. IEEE, 729-734.

Goyal, S., Damani, O. & Mahajan, A. 2021. Cost Optimization of Water Distribution Networks: Model Refinement Is Better Than Problem-Specific Solving Techniques. *arXiv preprint arXiv:2111.11865*.

Graham, S. & Marvin, S. 2002. *Splintering urbanism: networked infrastructures, technological mobilities and the urban condition*, Routledge.

Griggs, D., Nilsson, M., Stevance, A. & McCollum, D. 2017. *A guide to SDG interactions: from science to implementation*, International Council for Science, Paris.

Grover, A., Lall, S. & Maloney, W. 2022. *Place, productivity, and prosperity: Revisiting spatially targeted policies for regional development*, World Bank Publications.

Gudmundsson, A. & Mohajeri, N. 2013. Entropy and order in urban street networks. *Scientific Reports,* 3**,** 3324.

Gundogan, N. & Bicerli, M. K. 2009. Urbanization and labor market informality in developing countries.

Hagberg, A., Swart, P. & S Chult, D. 2008. Exploring network structure, dynamics, and function using NetworkX. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

Haghbayan, S. A., Geroliminis, N. & Akbarzadeh, M. 2021. Community detection in large scale congested urban road networks. *PLOS ONE,* 16**,** e0260201.

Halkos, G. & Gkampoura, E.-C. 2021. Where do we stand on the 17 Sustainable Development Goals? An overview on progress. *Economic Analysis and Policy,* 70**,** 94-122.

Han, B., Sun, D., Yu, X., Song, W. & Ding, L. 2020. Classification of Urban Street Networks Based on Tree-Like Network Features. *Sustainability,* 12**,** 628.

Hanberry, B. B. 2022. Imposing consistent global definitions of urban populations with gridded population density models: Irreconcilable differences at the national scale. *Landscape and Urban Planning,* 226**,** 104493.

Harenberg, S., Bello, G., Gjeltema, L., Ranshous, S., Harlalka, J., Seay, R., Padmanabhan, K. & Samatova, N. 2014. Community detection in large‐scale networks: a survey and empirical evaluation. *Wiley Interdisciplinary Reviews: Computational Statistics,* 6**,** 426-439.

Harris, L. M. 2021. Everyday Experiences of Water Insecurity: Insights from Underserved Areas of Accra, Ghana. *Daedalus,* 150**,** 64-84.

Harris, R. 2009. Slums. *In:* KITCHIN, R. & THRIFT, N. (eds.) *International Encyclopedia of Human Geography.* Oxford: Elsevier.

Harris, R. 2020. Neighbourhood upgrading: A fragmented global history. *Progress in Planning,* 142**,** 100432.

Hart, J. T. 1971. The inverse care law. *The Lancet,* 297**,** 405-412.

Hart, K. 1973. Informal Income Opportunities and Urban Employment in Ghana. *The Journal of Modern African Studies,* 11**,** 61-89.

Hennessey, D., Brooks, D., Fridman, A. & Breen, D. A simplification algorithm for visualizing the structure of complex graphs. 2008 12th International Conference Information Visualisation, 2008. IEEE, 616-625.

Henry, E., Bonnetain, L., Furno, A., Faouzi, N. E. E. & Zimeo, E. Spatio-temporal Correlations of Betweenness Centrality and Traffic Metrics. 2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), 5-7 June 2019 2019. 1-10.

Henson, R. M., Ortigoza, A., Martinez-Folgar, K., Baeza, F., Caiaffa, W., Vives Vergara, A., Diez Roux, A. V. & Lovasi, G. 2020. Evaluating the health effects of place-based slum upgrading physical environment interventions: A systematic review (2012–2018). *Social Science & Medicine,* 261**,** 113102.

Herrera, M., Abraham, E. & Stoianov, I. 2015. Graph-theoretic Surrogate Measures for Analysing the Resilience of Water Distribution Networks. *Procedia Engineering,* 119**,** 1241-1248.

Herrera, V. 2019. Reconciling global aspirations and local realities: Challenges facing the Sustainable Development Goals for water and sanitation. *World Development,* 118**,** 106-117.

Herschan, J., Tsinda, A., Okurut, K., Malcolm, R., Lapworth, D. J. & Pond, K. 2023. Progress of Using Risk Assessment to Manage Small Drinking-Water Supplies in Rwanda: A Preliminary Study. *Processes,* 11**,** 748.

Herslund, L. & Mguni, P. 2019. Examining urban water management practices – Challenges and possibilities for transitions to sustainable urban water management in Sub-Saharan cities. *Sustainable Cities and Society,* 48**,** 101573.

Heuvelink, G. B. M. 1998. Uncertainty analysis in environmental modelling under a change of spatial scale. *In:* FINKE, P. A., BOUMA, J. & HOOSBEEK, M. R. (eds.) *Soil and Water Quality at Different Scales: Proceedings of the Workshop "Soil and Water Quality at Different Scales" held 7–9 August 1996, Wageningen, The Netherlands.* Dordrecht: Springer Netherlands.

Hierink, F., Boo, G., Macharia, P. M., Ouma, P. O., Timoner, P., Levy, M., Tschirhart, K., Leyk, S., Oliphant, N., Tatem, A. J. & Ray, N. 2022. Differences between gridded population data impact measures of geographic access to healthcare in sub-Saharan Africa. *Communications Medicine,* 2**,** 117.

Hierink, F., Oladeji, O., Robins, A., Muñiz, M. F., Ayalew, Y. & Ray, N. 2023. A geospatial analysis of accessibility and availability to implement the primary healthcare roadmap in Ethiopia. *Communications Medicine,* 3**,** 140.

Higgins, C. D., Xi, Y. L., Widener, M., Palm, M., Vaughan, J., Miller, E. J., DeJohn, A. & Farber, S. 2022. Calculating place-based transit accessibility

Methods, tools and algorithmic dependence. *Journal of Transport and Land Use,* 15**,** 95-116.

Hill, P. H. 1952. Recent Slum Clearance and Urban Redevelopment Laws. *Wash. & Lee L. Rev.,* 9**,** 173.

Hillier, B. & Hanson, J. 1989. *The social logic of space*, Cambridge university press.

Hillier, B., Penn, A., Hanson, J., Grajewski, T. & Xu, J. 1993. Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: planning and design,* 20**,** 29-66.

Hitayezu, P., Rajashekar, A. & Stoelinga, D. 2018. The dynamics of unplanned settlements in the City of Kigali. *Laterite and International Gowth Center, Kigali*.

Ho, J. C., Russel, K. C. & Davis, J. 2013. The challenge of global water access monitoring: evaluating straight-line distance versus self-reported travel time among rural households in Mozambique. *Journal of Water and Health,* 12**,** 173-183.

Hodge, G. A. & Greve, C. 2017. On Public–Private Partnership Performance:A Contemporary Review. *Public Works Management & Policy,* 22**,** 55-78.

Hong, Y. & Yao, Y. 2019. Hierarchical community detection and functional area identification with OSM roads and complex graph theory. *International Journal of Geographical Information Science,* 33**,** 1569-1587.

Hopewell, M. R. & Graham, J. P. 2014. Trends in access to water supply and sanitation in 31 major sub-Saharan African cities: an analysis of DHS data from 2000 to 2012. *BMC Public Health,* 14**,** 208.

Hu, W., Li, X., Gao, X., Shang, T., Hou, P. & Li, D. A Vulnerability Identification Method of Power System Based on Infomap Community Discovery Algorithm. 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2), 22-24 Oct. 2021 2021. 2454-2459.

Huang, R. 2019. Simulating individual work trips for transit-facilitated accessibility study. *Environment and Planning B: Urban Analytics and City Science,* 46**,** 84-102.

Hudani, S. E. 2020. The Green Masterplan: Crisis, State Transition and Urban Transformation in Post-Genocide Rwanda. *International Journal of Urban and Regional Research,* 44**,** 673-690.

Huff, D. L. 1963. A Probabilistic Analysis of Shopping Center Trade Areas. *Land Economics,* 39**,** 81-90.

Hwang, H. & Lansey, K. 2017. Water Distribution System Classification Using System Characteristics and Graph-Theory Metrics. *Journal of Water Resources Planning and Management,* 143**,** 04017071.

Ioris, A. A. R. 2012. The geography of multiple scarcities: Urban development and water problems in Lima, Peru. *Geoforum,* 43**,** 612-622.

Isoke, J. & van Dijk, M. P. 2014. Factors influencing selection of drinking water technologies for urban informal settlements in Kampala. *Water and Environment Journal,* 28**,** 423-433.

Jaganyi, D., Njunwa, K., Nzayirambaho, M., Rutayisire, P. C., Manirakiza, V., Nsabimana, A. & Nduwayezu, G. 2018. Rwanda: National Urban policies and city profiles for Kigali and Huye. *79. glasgow: The GCRF centre for sustainable, healthy and learning cities and neighbourhoods (SHLC)*.

Jaglin, S. 2012. Networked Services and Features of African Urbanization: Other Path toward Globalization. *L'Espace géographique,* 41**,** 51-67.

Jang, K. M., Suh, H., Haddad, F. G., Sun, M., Duarte, F. & Kim, Y. 2024. Urban street clusters: unraveling the associations of street characteristics on urban vibrancy dynamics in age, time, and day. *Urban Informatics,* 3**,** 27.

Janoušková, S., Hák, T. & Moldan, B. 2018. Global SDGs Assessments: Helping or Confusing Indicators? *Sustainability,* 10**,** 1540.

Javed, M. A., Younis, M. S., Latif, S., Qadir, J. & Baig, A. 2018. Community detection in networks: A multidisciplinary review. *Journal of Network and Computer Applications,* 108**,** 87-111.

Jazayeri, P. & Moeini, R. 2024. District metered areas determination for water distribution networks using improved Girvan-Newman algorithm. *Ain Shams Engineering Journal,* 15**,** 102676.

Jeh, G. & Widom, J. 2002. SimRank: a measure of structural-context similarity. *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining.* Edmonton, Alberta, Canada: Association for Computing Machinery.

Jensen, H. A. & Jerez, D. J. 2018. A stochastic framework for reliability and sensitivity analysis of large scale water distribution networks. *Reliability Engineering & System Safety,* 176**,** 80-92.

Jha, N. 2023. The Impact of SDG-11 on Building Sustainable Cities: PPP Project Scenarios for Urban Infrastructure Development in India. *In:* DUTTA, V. & GHOSH, P. (eds.) *Sustainability: Science, Policy, and Practice in India: Challenges and Opportunities*. Cham: Springer International Publishing.

Jiang, B. & Claramunt, C. 2004. A Structural Approach to the Model Generalization of an Urban Street Network*. *GeoInformatica,* 8**,** 157-171.

Johnson, C. A., Flage, R. & Guikema, S. D. 2019. Characterising the robustness of coupled power-law networks. *Reliability Engineering & System Safety,* 191**,** 106560.

Jones, B. 2010. Slum Clearance, Privatization and Residualization: the Practices and Politics of Council Housing in Mid-twentieth-century England. *Twentieth Century British History,* 21**,** 510-539.

Jones, P. 2008. Different but the same? *City,* 12**,** 356-371.

Joshi, N., Gerlak, A. K., Hannah, C., Lopus, S., Krell, N. & Evans, T. 2023. Water insecurity, housing tenure, and the role of informal water services in Nairobi's slum settlements. *World Development,* 164**,** 106165.

Juju, D., Baffoe, G., Dam Lam, R., Karanja, A., Naidoo, M., Ahmed, A., Jarzebski, M. P., Saito, O., Fukushi, K., Takeuchi, K. & Gasparatos, A. 2020. Sustainability Challenges in Sub-Saharan Africa in the Context of the Sustainable Development Goals (SDGs). *In:* GASPARATOS, A., AHMED, A., NAIDOO, M., KARANJA, A., FUKUSHI, K., SAITO, O. & TAKEUCHI, K. (eds.) *Sustainability Challenges in Sub-Saharan Africa I: Continental Perspectives and Insights from Western and Central Africa.* Singapore: Springer Singapore.

Julien Chiquet, G. R., and Martina Sundqvist 2024. aricode: Efficient Computations of Standard Clustering Comparison Measures. version 1.0.3 ed.

Kabir, G., Tesfamariam, S., Hemsing, J. & Sadiq, R. 2020. Handling incomplete and missing data in water network database using imputation methods. *Sustainable and Resilient Infrastructure,* 5**,** 365-377.

Kaminsky, J. & Kumpel, E. 2018. Dry Pipes: Associations between Utility Performance and Intermittent Piped Water Supply in Low and Middle Income Countries. *Water,* 10**,** 1032.

Kang, C., Ma, X., Tong, D. & Liu, Y. 2012. Intra-urban human mobility patterns: An urban morphology perspective. *Physica A: Statistical Mechanics and its Applications,* 391**,** 1702-1717.

Kanuganti, S., Sarkar, A. K. & Singh, A. P. 2016. Evaluation of access to health care in rural areas using enhanced two-step floating catchment area (E2SFCA) method. *Journal of Transport Geography,* 56**,** 45-52.

Kao, T.-C. & Porter, M. A. 2018. Layer communities in multiplex networks. *Journal of Statistical Physics,* 173**,** 1286-1302.

Kapatsila, B., Palacios, M. S., Grisé, E. & El-Geneidy, A. 2023. Resolving the accessibility dilemma: Comparing cumulative and gravity-based measures of accessibility in eight Canadian cities. *Journal of Transport Geography,* 107**,** 103530.

Kawamoto, T. & Rosvall, M. 2015. Estimating the resolution limit of the map equation in community detection. *Physical Review E,* 91**,** 012809.

Kayaga, S., Fisher, J., Goodall, S., Kanyesigye, C., Kaggwa, R., Nambiro, M., Kitakufe, R., Otema, J. B., Mafunguro, R. & Ahabwe, G. 2020. Enhancing livelihoods of the urban poor through productive uses of utility-supplied water services – Evidence from Kampala, Uganda. *Cities,* 102**,** 102721.

Kaye-Essien, C. W. 2020. Understanding Absences and Presences of BID Policies: a Comparative Case of Accra and Cape Town. *Urban Forum,* 31**,** 177-195.

Kearns, A., Wright, V., Abrams, L. & Hazley, B. 2019. Slum clearance and relocation: a reassessment of social outcomes combining short-term and long-term perspectives. *Housing Studies,* 34**,** 201-225.

Keivani, R. & Werna, E. 2001. Modes of housing provision in developing countries. *Progress in Planning,* 55**,** 65-118.

Kelobonye, K., Zhou, H., McCarney, G. & Xia, J. 2020. Measuring the accessibility and spatial equity of urban services under competition using the cumulative opportunities measure. *Journal of Transport Geography,* 85**,** 102706.

Kempinska, K. & Murcio, R. 2019. Modelling urban networks using Variational Autoencoders. *Applied Network Science,* 4**,** 114.

Kenya National Bureau of Statistics (KNBS) 2019a. 2019 Kenya Population and Housing Census Volume II: distribution of population by administrative units Kenya National Bureau of Statistics Nairobi, Kenya.

Kenya National Bureau of Statistics (KNBS) 2019b. 2019 Kenya Population and Housing Census Volume IV: distribution of population by socio-economic characteristics. Kenya National Bureau of Statistics Nairobi, Kenya.

Kernighan, B. W. & Lin, S. 1970. An efficient heuristic procedure for partitioning graphs. *The Bell system technical journal,* 49**,** 291-307.

Kesmodel, U. S. 2018. Cross-sectional studies – what are they good for? *Acta Obstetricia et Gynecologica Scandinavica,* 97**,** 388-393.

Khaleghian, H. & Shan, Y. 2023. Developing a Data Quality Evaluation Framework for Sewer Inspection Data. *Water,* 15**,** 2043.

Khan, S. S., Te Lintelo, D. & Macgregor, H. 2023. Framing 'slums': global policy discourses and urban inequalities. *Environment and Urbanization,* 35**,** 74-90.

Khanani, R. S., Adugbila, E. J., Martinez, J. A. & Pfeffer, K. 2021. The Impact of Road Infrastructure Development Projects on Local Communities in Peri-Urban Areas: the Case of Kisumu, Kenya and Accra, Ghana. *International Journal of Community Well-Being,* 4**,** 33-53.

Khashoggi, B. F. & Murad, A. 2021. Use of 2SFCA Method to Identify and Analyze Spatial Access Disparities to Healthcare in Jeddah, Saudi Arabia. *Applied Sciences,* 11**,** 9537.

Khoa Bui, X., S. Marlim, M. & Kang, D. 2020. Water Network Partitioning into District Metered Areas: A State-Of-The-Art Review. *Water,* 12**,** 1002.

Kiani, B., Mohammadi, A., Bergquist, R. & Bagheri, N. 2021. Different configurations of the two-step floating catchment area method for measuring the spatial accessibility to hospitals for people living with disability: a cross-sectional study. *Archives of Public Health,* 79**,** 85.

Kim, S., Curran, K., Deng, L., Odhiambo, A., Oremo, J., Otieno, R., Omore, R., Handzel, T. & Quick, R. 2020. Backpack use as an alternative water transport method in Kisumu, Kenya. *Journal of Water, Sanitation and Hygiene for Development,* 10**,** 986-995.

Kirkley, A., Barbosa, H., Barthelemy, M. & Ghoshal, G. 2018. From the betweenness centrality in street networks to structural invariants in random planar graphs. *Nature Communications,* 9**,** 2501.

Kirkpatrick, C., Parker, D. & Zhang, Y.-F. 2006. An Empirical Analysis of State and Private-Sector Provision of Water Services in Africa. *The World Bank Economic Review,* 20**,** 143-163.

Kirstein, J. K., Klavs, H., Martin, R. & and Borup, M. 2019. A semi-automated approach to validation and error diagnostics of water network data. *Urban Water Journal,* 16**,** 1-10.

Kitchovitch, S. & Liò, P. 2011. Community Structure in Social Networks: Applications for Epidemiological Modelling. *PLOS ONE,* 6**,** e22220.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y. & Porter, M. A. 2014. Multilayer networks. *Journal of Complex Networks,* 2**,** 203-271.

Klar, B., Lee, J., Long, J. A. & Diab, E. 2023. The impacts of accessibility measure choice on public transit project evaluation: A comparative study of cumulative, gravity-based, and hybrid approaches. *Journal of Transport Geography,* 106**,** 103508.

Kleemeier, E. & Lockwood, H. 2012. Public-private partnerships for rural water services.

Kleiner, Y. & Rajani, B. 2001. Comprehensive review of structural deterioration of water mains: statistical models. *Urban Water,* 3**,** 131-150.

Klinkhamer, C., Krueger, E., Zhan, X., Blumensaat, F., Ukkusuri, S. & Rao, P. S. C. 2017. Functionally fractal urban networks: Geospatial co-location and homogeneity of infrastructure. *arXiv preprint arXiv:1712.03883*.

Klinkhamer, C., Zischg, J., Krueger, E., Yang, S., Blumensaat, F., Urich, C., Kaeseberg, T., Paik, K., Borchardt, D. & Silva, J. R. 2019. Topological convergence of urban infrastructure networks. *arXiv preprint arXiv:1902.01266*.

Knox, P. L. & Pacione, M. 1980. Locational behaviour, place preferences and the inverse care law in the distribution of primary medical care. *Geoforum,* 11**,** 43-55.

Kolowa, T. J., Daams, M. N. & Kuffer, M. 2024. Do informal settlements contribute to sprawl in Sub-Saharan African cities? *Sustainable Cities and Society,* 113**,** 105663.

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

Kranioti, A., Tsiotas, D. & Polyzos, S. 2022. The Topology of Cultural Destinations' Accessibility: The Case of Attica, Greece. *Sustainability,* 14**,** 1860.

KRB 2021. Road network in Kenya.

Krishna Reddy, P., Kitsuregawa, M., Sreekanth, P. & Rao, S. 2002. *A Graph Based Approach to Extract a Neighborhood Customer Community for Collaborative Filtering*.

Kropf, K. 2014. Ambiguity in the definition of built form. *Urban morphology,* 18**,** 41-57.

Krueger, E., Klinkhamer, C., Urich, C., Zhan, X. & Rao, P. S. C. 2017. Generic patterns in the evolution of urban water networks: Evidence from a large Asian city. *Physical Review E,* 95**,** 032312.

Kuc-Czarnecka, M., Markowicz, I. & Sompolska-Rzechuła, A. 2023. SDGs implementation, their synergies, and trade-offs in EU countries – Sensitivity analysis-based approach. *Ecological Indicators,* 146**,** 109888.

Kuffer, M., Orina, F., Sliuzas, R. & Taubenböck, H. Spatial patterns of slums: Comparing African and Asian cities. 2017 Joint Urban Remote Sensing Event (JURSE), 2017. IEEE, 1-4.

Kuffer, M., Owusu, M., Oliveira, L., Sliuzas, R. & van Rijn, F. 2022. The Missing Millions in Maps: Exploring Causes of Uncertainties in Global Gridded Population Datasets. *ISPRS International Journal of Geo-Information* [Online], 11.

Kuffer, M., Pfeffer, K. & Sliuzas, R. 2016. Slums from Space—15 Years of Slum Mapping Using Remote Sensing. *Remote Sensing,* 8**,** 455.

Kuffer, M., Wang, J., Nagenborg, M., Pfeffer, K., Kohli, D., Sliuzas, R. & Persello, C. 2018. The Scope of Earth-Observation to Improve the Consistency of the SDG Slum Indicator. *ISPRS International Journal of Geo-Information,* 7**,** 428.

Kut, E., Canan, F. T., Kahraman, E. D. & Cubukcu, K. M. 2016. Examining the relation between the urban pattern and urban history: Using graph theory-based network indices. *HISTORY URBANISM*.

Lagerberg, E. 2016. Visualization-Aided Design of Water Infrastructures in Informal Settlements in Nairobi, Kenya.

Lagunas, E., Marques, A. G., Chatzinotas, S. & Ottersten, B. Graph similarity based on graph Fourier distances. 2018 26th European Signal Processing Conference (EUSIPCO), 2018. IEEE, 877-881.

Lake Victoria South Water Works Development Agency (LVSWSB) 2021. Kisumu County Water Resource Masterplan Draft Report.

Lämmer, S., Gehlsen, B. & Helbing, D. 2006. Scaling laws in the spatial structure of urban road networks. *Physica A: Statistical Mechanics and its Applications,* 363**,** 89-95.

Lan, T., Zhang, H. & Li, Z. 2022. Exploring the evolution of road centrality: A case study of Hong Kong from 1976 to 2018. *Applied Geography,* 146**,** 102767.

Lancichinetti, A. & Fortunato, S. 2009. Community detection algorithms: a comparative analysis. *Physical review E,* 80**,** 056117.

Lancichinetti, A., Radicchi, F. & Ramasco, J. J. 2010. Statistical significance of communities in networks. *Physical Review E,* 81**,** 046110.

Lancichinetti, A., Radicchi, F., Ramasco, J. J. & Fortunato, S. 2011. Finding Statistically Significant Communities in Networks. *PLOS ONE,* 6**,** e18961.

Lara-Hernandez, J. A., Coulter, C. M. & Melis, A. 2020. Temporary appropriation and urban informality: Exploring the subtle distinction. *Cities,* 99**,** 102626.

Law, S., Shen, Y., Penn, A. & Karimi, K. Identifying street-character-weighted local area using locally weighted community detection methods the case study of London and Amsterdam. 12th International Space Syntax Symposium, SSS 2019, 2019. Space Syntax Symposium (SSS).

Lawhon, M. & Le Roux, L. 2019. Southern urbanism or a world of cities? Modes of enacting a more global urban geography in textbooks, teaching and research. *Urban Geography,* 40**,** 1251-1269.

Lawhon, M., Nilsson, D., Silver, J., Ernstson, H. & Lwasa, S. 2018. Thinking through heterogeneous infrastructure configurations. *Urban Studies,* 55**,** 720-732.

Le Blanc, D. 2015. Towards Integration at Last? The Sustainable Development Goals as a Network of Targets. *Sustainable Development,* 23**,** 176-187.

Letema, S., van Vliet, B. & van Lier, J. B. 2014. Sanitation policy and spatial planning in urban East Africa: Diverging sanitation spaces and actor arrangements in Kampala and Kisumu. *Cities,* 36**,** 1-9.

Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., Rose, A., Stevens, F. R., Blankespoor, B., Frye, C., Comenetz, J., Sorichetta, A., MacManus, K., Pistolesi, L., Levy, M., Tatem, A. J. & Pesaresi, M. 2019. The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth Syst. Sci. Data,* 11**,** 1385-1409.

Li, G., Xiong, Z. & Song, Y. A Systematic Review of Quantitative Measurement Methods for Accessibility of Urban Infrastructure. *In:* LI, J., LU, W., PENG, Y., YUAN, H. & WANG, D., eds. Proceedings of the 27th International Symposium on Advancement of Construction Management and Real Estate, 2023// 2023a Singapore. Springer Nature Singapore, 574-592.

Li, Z., Alakshendra, A. & Smith, S. 2023b. A People-Centered Perspective on Slum Formalization Policy. *Housing Policy Debate,* 33**,** 553-572.

Liang, H., Yan, Q., Yan, Y. & Zhang, Q. 2023. Using an improved 3SFCA method to assess inequities associated with multimodal accessibility to green spaces based on mismatches between supply and demand in the metropolitan of Shanghai, China. *Sustainable Cities and Society,* 91**,** 104456.

Liao, F. 2019. Space–time prism bounds of activity programs: a goal-directed search in multi-state supernetworks. *International Journal of Geographical Information Science,* 33**,** 900-921.

Lima, S., Brochado, A. & Marques, R. C. 2021. Public-private partnerships in the water sector: A review. *Utilities Policy,* 69**,** 101182.

Lin, W.-C. & Tsai, C.-F. 2020. Missing value imputation: a review and analysis of the literature (2006–2017). *Artificial Intelligence Review,* 53**,** 1487-1509.

Liu, L., Lyu, H., Zhao, Y. & Zhou, D. 2022. An Improved Two-Step Floating Catchment Area (2SFCA) Method for Measuring Spatial Accessibility to Elderly Care Facilities in Xi'an, China. *Int J Environ Res Public Health,* 19.

Liu, X., Cheng, H.-M. & Zhang, Z.-Y. 2019. Evaluation of community detection methods. *IEEE Transactions on Knowledge and Data Engineering,* 32**,** 1736-1746.

Liu, Y., Safavi, T., Dighe, A. & Koutra, D. 2018. Graph Summarization Methods and Applications: A Survey. *ACM Comput. Surv.,* 51**,** Article 62.

Liu, Y., Wang, H., Jiao, L., Liu, Y., He, J. & Ai, T. 2015. Road centrality and landscape spatial patterns in Wuhan Metropolitan Area, China. *Chinese Geographical Science,* 25**,** 511-522.

Liu, Y., Wei, X., Jiao, L. & Wang, H. 2016. Relationships between Street Centrality and Land Use Intensity in Wuhan, China. *Journal of Urban Planning and Development,* 142**,** 05015001.

Lorenz, I.-S., Pouls, K. & Pelz, P. F. Comparability of Water Infrastructure Resilience of Different Urban Structures. International Conference on Uncertainty in Mechanical Engineering, 2021. Springer, 81-90.

Lu, B., Sun, H., Harris, P., Xu, M. & Charlton, M. 2018. Shp2graph: Tools to Convert a Spatial Network into an Igraph Graph in R. *ISPRS International Journal of Geo-Information,* 7**,** 293.

Lucci, P., Bhatkal, T. & Khan, A. 2018. Are we underestimating urban poverty? *World Development,* 103**,** 297-310.

Luo, W. & Qi, Y. 2009. An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place,* 15**,** 1100-1107.

Luo, W. & Wang, F. 2003. Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and planning B: planning and design,* 30**,** 865-884.

Luo, W. & Whippo, T. 2012. Variable catchment sizes for the two-step floating catchment area (2SFCA) method. *Health & Place,* 18**,** 789-795.

Ma, D., He, F., Yue, Y., Guo, R., Zhao, T. & Wang, M. 2024. Graph convolutional networks for street network analysis with a case study of urban polycentricity in Chinese cities. *International Journal of Geographical Information Science,* 38**,** 931-955.

Ma, G., Ahmed, N. K., Willke, T. L. & Yu, P. S. 2021. Deep graph similarity learning: a survey. *Data Mining and Knowledge Discovery,* 35**,** 688-725.

Macharia, P. M., Moturi, A. K., Mumo, E., Giorgi, E., Okiro, E. A., Snow, R. W. & Ray, N. 2023. Modelling geographic access and school catchment areas across public primary schools to support subnational planning in Kenya. *Children's Geographies,* 21**,** 832-848.

Macharia, P. M., Mumo, E. & Okiro, E. A. 2021. Modelling geographical accessibility to urban centres in Kenya in 2019. *PLOS ONE,* 16**,** e0251624.

Mah, R. S. H. & Shacham, M. 1978. Pipeline Network Design and Synthesis. *In:* DREW, T. B., COKELET, G. R., HOOPES, J. W. & VERMEULEN, T. (eds.) *Advances in Chemical Engineering.* Academic Press.

Mahuve, F. E. & Tarimo, B. C. 2022. Integrating fuzzy set function into floating catchment area measures: a determination of spatial accessibility of service points. *Annals of GIS,* 28**,** 307-323.

Mair, M., Zischg, J., Rauch, W. & Sitzenfrei, R. 2017. Where to Find Water Pipes and Sewers?—On the Correlation of Infrastructure Networks in the Urban Environment. *Water,* 9**,** 146.

Mangioni, G., Jurman, G. & Domenico, M. D. 2020. Multilayer Flows in Molecular Networks Identify Biological Modules in the Human Proteome. *IEEE Transactions on Network Science and Engineering,* 7**,** 411-420.

Manirakiza, V., Mugabe, L., Nsabimana, A. & Nzayirambaho, M. 2019. City Profile: Kigali, Rwanda. *Environment and Urbanization ASIA,* 10**,** 290-307.

Maoulidi, M. 2010. A water and sanitation needs assessment for Kisumu city, Kenya.

Mapunda, D. W., Chen, S. S. & Yu, C. 2018. The role of informal small-scale water supply system in resolving drinking water shortages in peri-urban Dar Es Salaam, Tanzania. *Applied Geography,* 92**,** 112-122.

Marceau, D. J., Howarth, P. J. & Gratton, D. J. 1994. Remote sensing and the measurement of geographical entities in a forested environment. 1. The scale and spatial aggregation problem. *Remote Sensing of Environment,* 49**,** 93-104.

Marin, P. 2009a. *Public-Private Partnerships for Urban Water Utilities*.

Marin, P. 2009b. Public private partnerships for urban water utilities: A review of experiences in developing countries.

Marris, P. 1979. The meaning of slums and patterns of change. *International Journal of Urban and Regional Research,* 3**,** 419-441.

Marshall, S. 2016. Line structure representation for road network analysis. *Journal of Transport and Land Use,* 9**,** 29-64.

Marshall, S., Gil, J., Kropf, K., Tomko, M. & Figueiredo, L. 2018. Street Network Studies: from Networks to Models and their Representations. *Networks and Spatial Economics,* 18**,** 735-749.

Martínez-Santos, P. 2017. Determinants for water consumption from improved sources in rural villages of southern Mali. *Applied Geography,* 85**,** 113-125.

Marzouk, M., Hamid, S. A. & El-Said, M. 2015. A methodology for prioritizing water mains rehabilitation in Egypt. *HBRC Journal,* 11**,** 114-128.

Maschler, T. & Savic, D. A. 1999. Simplification of water supply network models through linearisation. *Cent. Water Syst. Rep,* 1**,** 119.

Masucci, A. P., Stanilov, K. & Batty, M. 2014. Exploring the evolution of London's street network in the information space: A dual approach. *Physical Review E,* 89**,** 012805.

Mata, A. S. d. 2020. Complex Networks: a Mini-review. *Brazilian Journal of Physics,* 50**,** 658-672.

Mavin, K. 1996. *Predicting the failure performance of individual water mains*, Urban Water Research Association of Australia.

Mazumder, R. K., Salman, A. M., Li, Y. & Yu, X. 2018. Performance Evaluation of Water Distribution Systems and Asset Management. *Journal of Infrastructure Systems,* 24**,** 03118001.

McCarthy, A. D., Chen, T. & Ebner, S. An exact no free lunch theorem for community detection. Complex Networks and Their Applications VIII: Volume 1 Proceedings of the Eighth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2019, 2019. Springer, 176-187.

McCartney, S. & Krishnamurthy, S. 2018. Neglected? Strengthening the Morphological Study of Informal Settlements. *Sage Open,* 8**,** 2158244018760375.

McDonald, D. A. 2018. Remunicipalization: The future of water services? *Geoforum,* 91**,** 47-56.

McGrail, M. R. & Humphreys, J. S. 2009. Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. *Applied Geography,* 29**,** 533-541.

Meijer, J. R., Huijbregts, M. A., Schotten, K. C. & Schipper, A. M. 2018. Global patterns of current and future road infrastructure. *Environmental Research Letters,* 13**,** 064006.

Meng, F., Fu, G., Farmani, R., Sweetapple, C. & Butler, D. 2018. Topological attributes of network resilience: A study in water distribution systems. *Water Research,* 143**,** 376-386.

Mengistie, B. M., Shi, W., Wong, M. S. & Zhu, R. 2023. Urban street network and data science based spatial connectivity evaluation of African cities: implications for sustainable urban development. *GeoJournal,* 88**,** 4753-4766.

Meredith, T. & MacDonald, M. 2017. Community-supported slum-upgrading: Innovations from Kibera, Nairobi, Kenya. *Habitat International,* 60**,** 1-9.

Merton, R. K. 1968. The Matthew Effect in Science. *Science,* 159**,** 56-63.

Miller, E. J. 2018. Accessibility: measurement and application in transportation planning. *Transport Reviews,* 38**,** 551-555.

Ministry of Lands Housing and Urban Development (MLHUD) 2008. National Slum Upgrading Strategy and Action Plan.

Ministry of Roads and Transport, T. R. o. K. 2023. Road Design Manual. NAIROBI.

Miranda, J., Scholz, I., Agard, J., Al Ghanim, K., Bobylev, S., Dube, O., Hathie, I., Kanie, N., Madise, N. & Malekpour, S. 2023. Times of crisis, times of change: Science for accelerating transformations to sustainable development. *United Nations, New York*.

Misselhorn, M. 2008. Position paper on informal settlement upgrading. *Part of a Strategy for the Second Economy for the Office of*.

Mittal, R. & Bhatia, M. P. S. 2021. Classification and Comparative Evaluation of Community Detection Algorithms. *Archives of Computational Methods in Engineering,* 28**,** 1417-1428.

Mizen, A., Fry, R., Grinnell, D. & Rodgers, S. E. 2015. Quantifying the Error Associated with Alternative GIS-based Techniques to Measure Access to Health Care Services. *AIMS Public Health,* 2**,** 746-761.

Mohajeri, N. & Gudmundsson, A. 2014. The Evolution and Complexity of Urban Street Networks. *Geographical Analysis,* 46**,** 345-367.

Mohajeri, N., Gudmundsson, A. & Scartezzini, J.-L. Expansion and densification of cities: linking urban form to ecology. Proceedings of International Conference CISBAT 2015 Future Buildings and Districts Sustainability from Nano to Urban Scale, 2015. LESO-PB, EPFL, 475-480.

Molden, D. 2020. Scarcity of water or scarcity of management? *International Journal of Water Resources Development,* 36**,** 258-268.

Momeni, A., Chauhan, V., Mahmoud, A. B., Piratla, K. R. & Safro, I. 2023. Generation of Synthetic Water Distribution Data Using a Multiscale Generator-Optimizer. *Journal of Pipeline Systems Engineering and Practice,* 14**,** 04022074.

Moreno, E. L. 2003. *Slums of the world: The face of urban poverty in the new millennium?: Monitoring the millennium development goal, target 11--world-wide slum dweller estimation*, Un-Habitat.

Moretto, L., Faldi, G., Ranzato, M., Rosati, F. N., Ilito Boozi, J.-P. & Teller, J. 2018. Challenges of water and sanitation service co-production in the global South. *Environment and Urbanization,* 30**,** 425-443.

Morzy, M., Kajdanowicz, T. & Szymański, B. K. 2016. Benford's Distribution in Complex Networks. *Scientific Reports,* 6**,** 34917.

Moyo, T. 2023. Subaltern urbanism as a way of life: Redefining and (re)imagining postcolonial urbanity in Bulawayo, Zimbabwe. *Journal of Urban Affairs,* 45**,** 428-440.

Muchadenyika, D. 2015. Slum upgrading and inclusive municipal governance in Harare, Zimbabwe: New perspectives for the urban poor. *Habitat International,* 48**,** 1-10.

Mugiraneza, T., Ban, Y. & Haas, J. 2019. Urban land cover dynamics and their impact on ecosystem services in Kigali, Rwanda using multi-temporal Landsat data. *Remote Sensing Applications: Society and Environment,* 13**,** 234-246.

Mukheibir, P. 2010. Water access, water scarcity, and climate change. *Environ Manage,* 45**,** 1027-39.

Müller, I., Taubenböck, H., Kuffer, M. & Wurm, M. 2020. Misperceptions of Predominant Slum Locations? Spatial Analysis of Slum Locations in Terms of Topography Based on Earth Observation Data. *Remote Sensing,* 12**,** 2474.

Munikoti, S., Lai, K. & Natarajan, B. 2021. Robustness assessment of Hetero-functional graph theory based model of interdependent urban utility networks. *Reliability Engineering & System Safety,* 212**,** 107627.

Munoz-Jofre, J., Hinojosa, S., Mascle-Allemand, A. L. & Temprano, J. 2023. A selectivity index for public-private partnership projects in the urban water and sanitation sector in Latin America and the caribbean. *J Environ Manage,* 335**,** 117564.

Mutikanga, H. E., Sharma, S. & Vairavamoorthy, K. 2009. Water loss management in developing countries: Challenges and prospec. *Journal AWWA,* 101**,** 57-68.

Mvulirwenande, S., Wehn, U. & Alaerts, G. 2019. Policy factors explaining the failure of delegated management in water supply: evidence from Ghana. *Water International,* 44**,** 14-30.

Najlis, P. & Edwards, A. 1991. The International Drinking Water Supply and Sanitation Decade in retrospect and implications for the future. *Natural Resources Forum,* 15**,** 110-117.

Nallathiga, R. 2012. Slum redevelopment in cities: current approaches and alternate models. *Insititute of Town Planners, India Journal***,** 9-2.

NASA & METI 2011. ASTER Global Digital Elevation Model.

National Institute of Statistics of Rwanda (NISR) 2018. The Fifth Integrated Household Living Conditions Survey: Main Indicators Report.

National Institute of Statistics of Rwanda (NISR) 2021. Rwanda Demographic and Health Survey 2019-20.

National Institute of Statistics of Rwanda (NISR) 2023a. Main Indicators: 5th Rwanda Population and Housing Census (PHC).

National Institute of Statistics of Rwanda (NISR) 2023b. Thematic Report:Population Size, Structure and Spatial Distribution.

Nauges, C. & Whittington, D. 2010. Estimation of Water Demand in Developing Countries: An Overview. *The World Bank Research Observer,* 25**,** 263-294.

Nduwayezu, G., Manirakiza, V., Mugabe, L. & Malonza, J. M. 2021. Urban Growth and Land Use/Land Cover Changes in the Post-Genocide Period, Kigali, Rwanda. *Environment and Urbanization ASIA,* 12**,** S127-S146.

Nduwayezu, G., Sliuzas, R. & Kuffer, M. 2016. Modeling urban growth in Kigali city Rwanda. *Rwanda journal,* 1.

Neal, Z. 2018. Is the Urban World Small? The Evidence for Small World Structure in Urban Networks. *Networks and Spatial Economics,* 18**,** 615-631.

Neal, Z., Derudder, B. & Liu, X. 2021. Using urban networks to gain new insight into old questions: Community, economy, bureaucracy. *Journal of Urban Affairs,* 43**,** 2-15.

Negadi, M. A., Mebirouk, H. & Djedid, A. 2023. Exploring Urban Coherence through Fractality, Connectivity, and Arteriality of the Urban Street Network: Comparative Study of Five Medium-Sized Algerian Cities. *Journal of Urban Planning and Development,* 149**,** 04023038.

Nel, M., Simuyaba, M., Muchelenje, J., Chirwa, T., Simwinga, M., Speight, V., Mhlanga, Z., Jacobs, H., Nel, N., Seeley, J., Mwanaumo, E., Viljoen, L., Hoddinott, G. & Bond, V. 2023. Broad Brush Surveys: a rapid qualitative assessment approach for water and sanitation infrastructure in urban sub-Saharan cities. *Frontiers in Sustainable Cities,* 5.

Neutens, T., Delafontaine, M., Scott, D. M. & De Maeyer, P. 2012. An analysis of day-to-day variations in individual space–time accessibility. *Journal of Transport Geography,* 23**,** 81-91.

Newman, D. A. 2014. Missing Data:Five Practical Guidelines. *Organizational Research Methods,* 17**,** 372-411.

Newman, M. E. & Girvan, M. 2004. Finding and evaluating community structure in networks. *Physical review E,* 69**,** 026113.

Ngben, J. M. & Yakubu, I. 2023. 'You just have to stay awake and keep spying on the pipes': Everyday experiences of piped water access in Tamale, Ghana. *Habitat International,* 138**,** 102865.

Nijman, J. 2007. Introduction—Comparative Urbanism. *Urban Geography,* 28**,** 1-6.

Nijman, J. 2008. Against the odds: Slum rehabilitation in neoliberal Mumbai. *Cities,* 25**,** 73-85.

Nilsson, D. 2011. *Pipes, progress, and poverty: Social and technological change in urban water provision in Kenya and Uganda 1895-2010.* KTH Royal Institute of Technology.

Nilsson, D. & Nyanchaga, E. N. 2008. Pipes and politics: a century of change and continuity in Kenyan urban water supply. *The Journal of Modern African Studies,* 46**,** 133-158.

Nilsson, M., Griggs, D. & Visbeck, M. 2016. Policy: Map the interactions between Sustainable Development Goals. *Nature,* 534**,** 320-322.

Nolan, L. B. 2015. Slum Definitions in Urban India: Implications for the Measurement of Health Inequalities. *Popul Dev Rev,* 41**,** 59-84.

Nor, M. M., Noor, N. M. & Asmawi, M. Z. 2021. Bibliometric Analysis on Street Network components in Influencing Genome of Urban Morphologies. *Journal of Architecture, Planning and Construction Management (JAPCM),* 11.

Nyika, J. & Dinka, M. O. 2023. Water Challenges in Urban Sub-Saharan Africa. *In:* NYIKA, J. & DINKA, M. O. (eds.) *Water Challenges in Rural and Urban Sub-Saharan Africa and their Management.* Cham: Springer Nature Switzerland.

Nzengya, D. M. 2015. Exploring the challenges and opportunities for master operators and water kiosks under Delegated Management Model (DMM): A study in Lake Victoria region, Kenya. *Cities,* 46**,** 35-43.

Nzengya, D. M. 2018. Improving water service to the urban poor through delegated management: Lessons from the city of Kisumu, Kenya. *Development Policy Review,* 36**,** 190-202.

O'Sullivan, D. 2014. Spatial Network Analysis. *In:* FISCHER, M. M. & NIJKAMP, P. (eds.) *Handbook of Regional Science.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Oak Ridge National Laboratory 2021. Population count.

Ocholla, G., Letema, S. & Mireri, C. 2022. Socioeconomic determinants of water delivery satisfaction in a medium sub-Saharan Africa city: A case of Kisumu, Kenya. *Water Supply,* 22**,** 8682-8697.

OECD n.d. Principles for Public Governance of Public-Private Partnerships - OECD.

Okotto-Okotto, J., Okotto, L., Price, H., Pedley, S. & Wright, J. 2015. A Longitudinal Study of Long-Term Change in Contamination Hazards and Shallow Well Quality in Two Neighbourhoods of Kisumu, Kenya. *International Journal of Environmental Research and Public Health,* 12**,** 4275-4291.

Okotto, L., Okotto-Okotto, J., Price, H., Pedley, S. & Wright, J. 2015. Socio-economic aspects of domestic groundwater consumption, vending and use in Kisumu, Kenya. *Applied Geography,* 58**,** 189-197.

Okpala, D. C. I. 1992. Housing production systems and technologies in developing countries: a review of the experiences and possible future trends/prospects. *Habitat International,* 16**,** 9-32.

Okyere, S. A., Diko, S. K., Hiraoka, M. & Kita, M. 2017. An Urban "Mixity": Spatial Dynamics of Social Interactions and Human Behaviors in the Abese informal Quarter of La Dadekotopon, Ghana. *Urban Science,* 1**,** 13.

Olthuis, K., Benni, J., Eichwede, K. & Zevenbergen, C. 2015. Slum Upgrading: Assessing the importance of location and a plea for a spatial approach. *Habitat International,* 50**,** 270-288.

Olu Sule, R. A. 1990. Recent slum clearance exercise in Lagos (Nigeria): victims or beneficiaries? *GeoJournal,* 22**,** 81-91.

ONAPO 1991. [The population of Rwanda as of January 1, 1991]. *Imbonezamuryango***,** 27-30.

Ondigo, D. A., Kavoo, A. M. & Kebwaro, J. 2018. Water Resources and Management under Increasing Urban Demography: A Kenyan Perspective—A Review. *Journal of Water Resource and Protection,* 10**,** 919-938.

OpenStreetMap contributors 2023. Planet dump retrieved from https://planet.osm.org.

Opryszko, M. C., Guo, Y., MacDonald, L., MacDonald, L., Kiihl, S. & Schwab, K. J. 2013. Impact of water-vending kiosks and hygiene education on household drinking water quality in rural Ghana. *Am J Trop Med Hyg,* 88**,** 651-60.

Ortega, A., Frossard, P., Kovačević, J., Moura, J. M. & Vandergheynst, P. 2018. Graph signal processing: Overview, challenges, and applications. *Proceedings of the IEEE,* 106**,** 808-828.

Osman, M. S., Abu-Mahfouz, A. M. & Page, P. R. 2018. A survey on data imputation techniques: Water distribution system as a use case. *IEEE Access,* 6**,** 63279-63291.

Othoo, C. O., Dulo, S. O., Olago, D. O. & Ayah, R. 2020. Proximity density assessment and characterization of water and sanitation facilities in the informal settlements of Kisumu city: Implications for public health planning. *Journal of UOEH,* 42**,** 237-249.

Otiso, K. M. 2003. State, voluntary and private sector partnerships for slum upgrading and basic service delivery in Nairobi City, Kenya. *Cities,* 20**,** 221-229.

Ozanne, H. S. 2011. Chapter 2 - Route Selection. *In:* MENON, E. S. (ed.) *Pipeline Planning and Construction Field Manual.* Boston: Gulf Professional Publishing.

Paez, D. & Filion, Y. 2017. Generation and Validation of Synthetic WDS Case Studies Using Graph Theory and Reliability Indexes. *Procedia Engineering,* 186**,** 143-151.

Pagano, A., Sweetapple, C., Farmani, R., Giordano, R. & Butler, D. 2019. Water Distribution Networks Resilience Analysis: a Comparison between Graph Theory-Based Approaches and Global Resilience Analysis. *Water Resources Management,* 33**,** 2925-2940.

Palacios-Lopez, D., Bachofer, F., Esch, T., Heldens, W., Hirner, A., Marconcini, M., Sorichetta, A., Zeidler, J., Kuenzer, C., Dech, S., Tatem, A. J. & Reinartz, P. 2019. New Perspectives for Mapping Global Population Distribution Using World Settlement Footprint Products. *Sustainability,* 11**,** 6056.

Panday, P. K. 2020. Urban Slum Upgrading Best Practices in Bangladesh. *In:* PANDAY, P. K. (ed.) *The Face of Urbanization and Urban Poverty in Bangladesh: Explaining the Slum Development Initiatives in the light of Global Experiences.* Singapore: Springer Singapore.

Parida, D. & Agrawal, S. 2023. Southern urbanism: a systematic review of concepts, debates, and future directions. *GeoJournal,* 88**,** 2587-2608.

Patel, A. & Killemsetty, N. A Systematic Review of Slum Policies in the Global South: What Could We Learn from Successes and Failures of the Past? Institute for Equity and Resilience: Annual World Bank Conference on Land and Poverty, Washington DC, 2020.

Patel, S. 2013. Upgrade, rehouse or resettle? An assessment of the Indian government's Basic Services for the Urban Poor (BSUP) programme. *Environment & Urbanization,* 25**,** 177-188.

Pebesma, E. & Bivand, R. 2023. *Spatial data science: With applications in R*, Chapman and Hall/CRC.

Peel, L., Larremore, D. B. & Clauset, A. 2017. The ground truth about metadata and community detection in networks. *Science Advances,* 3**,** e1602548.

Perelman, L. & Ostfeld, A. 2011. Topological clustering for water distribution systems analysis. *Environmental Modelling & Software,* 26**,** 969-972.

Pickering, A. J. & Davis, J. 2012. Freshwater Availability and Water Fetching Distance Affect Child Health in Sub-Saharan Africa. *Environmental Science & Technology,* 46**,** 2391-2397.

Pierce, G. 2017. Why is basic service access worse in slums? A synthesis of obstacles. *Development in Practice,* 27**,** 288-300.

Pieterse, E. 2011. Grasping the unknowable: coming to grips with African urbanisms. *Social Dynamics,* 37**,** 5-23.

Pons, P. & Latapy, M. Computing communities in large networks using random walks. Computer and Information Sciences-ISCIS 2005: 20th International Symposium, Istanbul, Turkey, October 26-28, 2005. Proceedings 20, 2005. Springer, 284-293.

Poorthuis, A. 2018. How to Draw a Neighborhood? The Potential of Big Data, Regionalization, and Community Detection for Understanding the Heterogeneous Nature of Urban Neighborhoods. *Geographical Analysis,* 50**,** 182-203.

Porse, E. & Lund, J. 2016. Network Analysis and Visualizations of Water Resources Infrastructure in California: Linking Connectivity and Resilience. *Journal of Water Resources Planning and Management,* 142**,** 04015041.

Porta, S., Crucitti, P. & Latora, V. 2006. The network analysis of urban streets: A dual approach. *Physica A: Statistical Mechanics and its Applications,* 369**,** 853-866.

Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Cardillo, A., Belli, E., Càrdenas, F., Cormenzana, B. & Latora, L. 2012. Street Centrality and the Location of Economic Activities in Barcelona. *Urban Studies,* 49**,** 1471-1488.

Portes, A. 1971. The urban slum in Chile: types and correlates. *Land Economics,* 47**,** 235-248.

Posada, H. M. & Moreno-Monroy, A. I. 2017. Informality, city structure and rural–urban migration in Latin America. *The Annals of Regional Science,* 59**,** 345-369.

Post, A. & Ray, I. 2020. Hybrid Modes of Urban Water Delivery in Low- and Middle-Income Countries. Oxford University Press.

Post, A. E. 2022. Informality and Politics in the Global South: Three Perspectives. *Perspectives on Politics,* 20**,** 1386-1390.

Pot, F. J., van Wee, B. & Tillema, T. 2021. Perceived accessibility: What it is and why it differs from calculated accessibility measures based on spatial data. *Journal of Transport Geography,* 94**,** 103090.

PPP Authority 2008. Public-Private Partnership Handbook. Dhaka, Bangladesh.

Price, H. D., Okotto, L. G., Okotto-Okotto, J., Pedley, S. & Wright, J. 2018. A participatory methodology for future scenario analysis of sub-national water and sanitation access: case study of Kisumu, Kenya. *Water International,* 43**,** 591-602.

Prieto-Curiel, R., Patino, J. E. & Anderson, B. 2023. Scaling of the morphology of African cities. *Proceedings of the National Academy of Sciences,* 120**,** e2214254120.

Prieto-Curiel, R., Schumann, A., Heo, I. & Heinrigs, P. 2022. Detecting cities with high intermediacy in the African urban network. *Computers, Environment and Urban Systems,* 98**,** 101869.

Pueyo, O., Pueyo, X. & Patow, G. 2019. An overview of generalization techniques for street networks. *Graphical Models,* 106**,** 101049.

Pumain, D. & Guerois, M. 2004. Scaling laws and urban systems. *Santa Fe Institute, Working Paper***,** 002.

Pung, J., D'Souza, R. M., Ghosal, D. & Zhang, M. 2022. A road network simplification algorithm that preserves topological properties. *Applied Network Science,* 7**,** 79.

Punmia, B., Jain, A. K. & Jain, A. K. 1995. *Water supply engineering*, Firewall Media.

Ray, N. & Ebener, S. 2008. AccessMod 3.0: computing geographic coverage and accessibility to health care services using anisotropic movement of patients. *International Journal of Health Geographics,* 7**,** 63.

Rebelo, M. & Matos, J. S. 2022. Urban Water Deficit in Sub-Saharan African Cities. *In:* LEAL FILHO, W., AZUL, A. M., BRANDLI, L., LANGE SALVIA, A. & WALL, T. (eds.) *Clean Water and Sanitation.* Cham: Springer International Publishing.

Recio, R. B., Mateo-Babiano, I. & Roitman, S. 2017. Revisiting policy epistemologies on urban informality: Towards a post-dualist view. *Cities,* 61**,** 136-143.

Revi, A. 2016. Afterwards: Habitat III and the Sustainable Development Goals. *Urbanisation,* 1**,** x-xiv.

Reza, S., Ferreira, M. C., Machado, J. J. M. & Tavares, J. M. R. S. 2024. Road networks structure analysis: A preliminary network science-based approach. *Annals of Mathematics and Artificial Intelligence,* 92**,** 215-234.

Richmond, A., Myers, I. & Namuli, H. 2018. Urban Informality and Vulnerability: A Case Study in Kampala, Uganda. *Urban Science,* 2**,** 22.

Riley, L. W., Ko, A. I., Unger, A. & Reis, M. G. 2007. Slum health: Diseases of neglected populations. *BMC International Health and Human Rights,* 7**,** 2.

Robinson, J. 2016. Comparative Urbanism: New Geographies and Cultures of Theorizing the Urban. *International Journal of Urban and Regional Research,* 40**,** 187-199.

Roeger, A. & Tavares, A. F. 2018. Water safety plans by utilities: A review of research on implementation. *Utilities Policy,* 53**,** 15-24.

Rosvall, M., Axelsson, D. & Bergstrom, C. T. 2009. The map equation. *The European Physical Journal Special Topics,* 178**,** 13-23.

Rosvall, M., Delvenne, J. C., Schaub, M. T. & Lambiotte, R. 2019. Different approaches to community detection. *Advances in network clustering and blockmodeling***,** 105-119.

Roy, A. 2011. Slumdog Cities: Rethinking Subaltern Urbanism. *International Journal of Urban and Regional Research,* 35**,** 223-238.

Roy, A. 2014. Slum-free cities of the Asian century: Postcolonial government and the project of inclusive growth. *Singapore Journal of Tropical Geography,* 35**,** 136-150.

RTDA 2022. Rwanda Classified Road Network

Ruan, N., Jin, R. & Huang, Y. Distance preserving graph simplification. 2011 IEEE 11th International Conference on Data Mining, 2011. IEEE, 1200-1205.

Rui, Y. & Ban, Y. 2014. Exploring the relationship between street centrality and land use in Stockholm. *International Journal of Geographical Information Science,* 28**,** 1425-1438.

Ruiz-Villaverde, A., Chica-Olmo, J. & González-Gómez, F. 2018. Do small municipalities imitate larger ones? Diffusion of water privatization policies. *Urban Water Journal,* 15**,** 138-149.

Rusca, M., Alda-Vidal, C., Hordijk, M. & Kral, N. 2017. Bathing without water, and other stories of everyday hygiene practices and risk perception in urban low-income areas: the case of Lilongwe, Malawi. *Environment and Urbanization,* 29**,** 533-550.

Rusca, M. & Cleaver, F. 2022. Unpacking everyday urbanism: Practices and the making of (un)even urban waterscapes. *WIREs Water,* 9**,** e1581.

Rwandan Ministry of Infrastructure (MININFRA) 2015a. National informal settlement upgrading strategy.

Rwandan Ministry of Infrastructure (MININFRA) 2015b. National Urbanisation Policy.

Sambu, D. 2016. Impact of global initiatives on drinking water access in Africa. *African Geographical Review,* 35**,** 151-167.

Sambu, D. K. & Tarhule, A. 2013. Progress of water service providers in meeting millennium development goals in Kenya. *African Geographical Review,* 32**,** 105-124.

Santonastaso, G. F., Di Nardo, A., Creaco, E., Musmarra, D. & Greco, R. 2021. Comparison of topological, empirical and optimization-based approaches for locating quality detection points in water distribution networks. *Environmental Science and Pollution Research,* 28**,** 33844-33853.

Sarbu, I. & Ostafe, G. 2016. Optimal design of urban water supply pipe networks. *Urban Water Journal,* 13**,** 521-535.

Sarkar, A. 2022. Access to Affordable, Safe Water Supply in Urban Kenya: A Macro Analysis. *In:* SARKAR, A. (ed.) *Water Insecurity and Water Governance in Urban Kenya: Policy and Practice.* Cham: Springer International Publishing.

Scheer, B. C. 2015. The epistemology of urban morphology. *Urban Morphology,* 20**,** 5-17.

Schiavina, M., Freire, S., Carioli, A. & MacManus, K. 2020. GHS-POP R2023A - GHS population grid multitemporal (1975–2030). *In:* EUROPEAN COMMISSION, J. R. C. J. (ed.).

Schieber, T. A., Carpi, L., Díaz-Guilera, A., Pardalos, P. M., Masoller, C. & Ravetti, M. G. 2017. Quantification of network structural dissimilarities. *Nature Communications,* 8**,** 13928.

Schindler, S. 2017. Towards a paradigm of Southern urbanism. *City,* 21**,** 47-64.

Schipper, A. M., Hilbers, J. P., Meijer, J. R., Antão, L. H., Benítez-López, A., de Jonge, M. M. J., Leemans, L. H., Scheper, E., Alkemade, R., Doelman, J. C., Mylius, S., Stehfest, E., van Vuuren, D. P., van Zeist, W.-J. & Huijbregts, M. A. J. 2020. Projecting terrestrial biodiversity intactness with GLOBIO 4. *Global Change Biology,* 26**,** 760-771.

Schläpfer, M., Bettencourt, L. M. A., Grauwin, S., Raschke, M., Claxton, R., Smoreda, Z., West, G. B. & Ratti, C. 2014. The scaling of human interactions with city size. *Journal of The Royal Society Interface,* 11**,** 20130789.

Schlosser, C. A., Strzepek, K., Gao, X., Fant, C., Blanc, É., Paltsev, S., Jacoby, H., Reilly, J. & Gueneau, A. 2014. The future of global water stress: An integrated assessment. *Earth's Future,* 2**,** 341-361.

Schnettler, S. 2009. A structured overview of 50 years of small-world research. *Social Networks,* 31**,** 165-178.

Schwartz, K. & Boakye-Ansah, A. 2023. Pragmatism as an approach for decision-making: Why two Kenyan water utilities opted for pre-paid water dispensers. *Utilities Policy,* 84**,** 101623.

Schwartz, K. & Sanga, A. 2010. Partnerships between utilities and small-scale providers: Delegated management in Kisumu, Kenya. *Physics and Chemistry of the Earth, Parts A/B/C,* 35**,** 765-771.

Scibetta, M., Boano, F., Revelli, R. & Ridolfi, L. 2013. Community detection as a tool for complex pipe network clustering. *Europhysics Letters,* 103**,** 48001.

Serra, M., Gil, J. & Pinho, P. 2016. Towards an understanding of morphogenesis in metropolitan street-networks. *Environment and Planning B: Urban Analytics and City Science,* 44**,** 272-293.

Serra, M. & Hillier, B. 2019. Angular and Metric Distance in Road Network Analysis: A nationwide correlation study. *Computers, Environment and Urban Systems,* 74**,** 194-207.

Seyfried, M. S. & Wilcox, B. P. 1995. Scale and the Nature of Spatial Variability: Field Examples Having Implications for Hydrologic Modeling. *Water Resources Research,* 31**,** 173-184.

Shang, W.-L., Chen, Y., Bi, H., Zhang, H., Ma, C. & Ochieng, W. Y. 2020. Statistical Characteristics and Community Analysis of Urban Road Networks. *Complexity,* 2020**,** 6025821.

Shao, Y. & Luo, W. 2022. Supply-demand adjusted two-steps floating catchment area (SDA-2SFCA) model for measuring spatial access to health care. *Social Science & Medicine,* 296**,** 114727.

Sharma, A. N., Dongre, S. R., Gupta, R. & Ormsbee, L. 2022a. Multiphase Procedure for Identifying District Metered Areas in Water Distribution Networks Using Community Detection, NSGA-III Optimization, and Multiple Attribute Decision Making. *Journal of Water Resources Planning and Management,* 148**,** 04022040.

Sharma, A. N., Dongre, S. R., Gupta, R., Pandey, P. & Bokde, N. D. 2022b. Partitioning of water distribution network into district metered areas using existing valves. *Computer Modeling in Engineering & Sciences,* 131**,** 1515-1537.

Sheppard, E., Leitner, H. & Maringanti, A. 2013. Provincializing Global Urbanism: A Manifesto. *Urban Geography,* 34**,** 893-900.

Sheppard, E., Sparks, T. & Leitner, H. 2020. World Class Aspirations, Urban Informality, and Poverty Politics: A North–South Comparison. *Antipode,* 52**,** 393-407.

Shi, X., Liu, D. & Gan, J. 2024. A Study on the Relationship between Road Network Centrality and the Spatial Distribution of Commercial Facilities—A Case of Changchun, China. *Sustainability,* 16**,** 3920.

Shields, K. F., Bain, R. E. S., Cronk, R., Wright, J. A. & Bartram, J. 2015. Association of Supply Type with Fecal Contamination of Source Water and Household Stored Drinking Water in Developing Countries: A Bivariate Meta-analysis. *Environmental Health Perspectives,* 123**,** 1222-1231.

Shimomura, L. C., Oyamada, R. S., Vieira, M. R. & Kaster, D. S. 2021. A survey on graph-based methods for similarity searches in metric spaces. *Information Systems,* 95**,** 101507.

Shuang, Q., Liu, H. J. & Porse, E. 2019. Review of the Quantitative Resilience Methods in Water Distribution Networks. *Water,* 11**,** 1189.

Shushu, U. P., Komakech, H. C., Dodoo-Arhin, D., Ferras, D. & Kansal, M. L. 2021. Managing non-revenue water in Mwanza, Tanzania: A fast-growing sub-Saharan African city. *Scientific African,* 12**,** e00830.

Siddiq, F. & Taylor, B. 2021. Tools of the trade? Assessing the progress of accessibility measures for planning practice. *Journal of the American Planning Association,* 87**,** 497-511.

Signorelli, M. & Cutillo, L. 2022. On community structure validation in real networks. *Computational Statistics,* 37**,** 1165-1183.

Signorelli, M., Vinciotti, V. & Wit, E. C. 2016. NEAT: an efficient network enrichment analysis test. *BMC Bioinformatics,* 17**,** 352.

Sima, L. C., Kelner-Levine, E., Eckelman, M. J., McCarty, K. M. & Elimelech, M. 2013. Water flows, energy demand, and market analysis of the informal water sector in Kisumu, Kenya. *Ecological Economics,* 87**,** 137-144.

Sinharoy, S. S., Pittluck, R. & Clasen, T. 2019. Review of drivers and barriers of water and sanitation policies for urban informal settlements in low-income and middle-income countries. *Utilities Policy,* 60**,** 100957.

Smiley, S. L. 2020. Heterogeneous water provision in Dar es Salaam: The role of networked infrastructures and alternative systems in informal areas. *Environment and Planning E: Nature and Space,* 3**,** 1215-1231.

Smiljanić, J., Blöcker, C., Holmgren, A., Edler, D., Neuman, M. & Rosvall, M. 2023. Community detection with the map equation and infomap: Theory and applications. *arXiv preprint arXiv:2311.04036*.

Smit, S., Musango, J. K., Kovacic, Z. & Brent, A. C. 2017. Conceptualising slum in an urban African context. *Cities,* 62**,** 107-119.

Smith, N. R., Zivich, P. N., Frerichs, L. M., Moody, J. & Aiello, A. E. 2020. A Guide for Choosing Community Detection Algorithms in Social Network Studies: The Question Alignment Approach. *American Journal of Preventive Medicine,* 59**,** 597-605.

Soldi, D., Candelieri, A. & Archetti, F. 2015. Resilience and Vulnerability in Urban Water Distribution Networks through Network Theory and Hydraulic Simulation. *Procedia Engineering,* 119**,** 1259-1268.

Soliman, A. M. 2020. Understanding sustainability transitions for urban informality in the Middle East. *International Journal of Urban Sustainable Development,* 12**,** 340-359.

Solomakhina, N., Hubauer, T. & Becher, S. On improving data quality and topology in vector spatial data. 2016.

Sorensen, A. 2018. Institutions and Urban Space: Land, Infrastructure, and Governance in the Production of Urban Property. *Planning Theory & Practice,* 19**,** 21-38.

Stacherl, B. & Sauzet, O. 2023. Gravity models for potential spatial healthcare access measurement: a systematic methodological review. *International Journal of Health Geographics,* 22**,** 34.

Stavroulaki, I., Marcus, L., Pont, M. B. & Nilsson, L. C. S. Representations of street networks in space syntax towards flexible maps and multiple graphs. 11th International Space Syntax Symposium, SSS 2017, Lisbon, Portugal, 3-7 July 2017, 2017. 174.1-174.16.

Stępniak, M. & Jacobs-Crisioni, C. 2017. Reducing the uncertainty induced by spatial aggregation in accessibility and spatial interaction applications. *Journal of Transport Geography,* 61**,** 17-29.

Stevens, F. R., Gaughan, A. E., Linard, C. & Tatem, A. J. 2015. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLOS ONE,* 10**,** e0107042.

Stewart, J. 2005. A review of UK housing policy: ideology and public health. *Public Health,* 119**,** 525-534.

Steyn, G. 2012. Modernity and Informality - Conflicting Forces Shaping Urban Form in Kisumu, Kenya.

Strano, E., Nicosia, V., Latora, V., Porta, S. & Barthélemy, M. 2012. Elementary processes governing the evolution of road networks. *Scientific Reports,* 2**,** 296.

Šuba, R., Meijers, M. & Oosterom, P. V. 2016. Continuous Road Network Generalization throughout All Scales. *ISPRS International Journal of Geo-Information,* 5**,** 145.

Subal, J., Paal, P. & Krisp, J. M. 2021. Quantifying spatial accessibility of general practitioners by applying a modified huff three-step floating catchment area (MH3SFCA) method. *International Journal of Health Geographics,* 20**,** 9.

Šubelj, L. 2018. Convex skeletons of complex networks. *Journal of The Royal Society Interface,* 15**,** 20180422.

Sulem, D., Kenlay, H., Cucuringu, M. & Dong, X. 2024. Graph similarity learning for change-point detection in dynamic networks. *Machine Learning,* 113**,** 1-44.

Sultana, F. 2020. Embodied Intersectionalities of Urban Citizenship: Water, Infrastructure, and Gender in the Global South. *Annals of the American Association of Geographers,* 110**,** 1407-1424.

Svensson, J., Dehn, J. & Reinikka, R. 2003. *Cities Alliance Annual Report 2003*.

Tan, J., Wang, X. & Pan, J. 2021. The effect of population distribution measures on evaluating spatial accessibility of primary health-care institutions: A case study from China. *Geospatial Health,* 16.

Tao, Z., Cheng, Y. & Liu, J. 2020. Hierarchical two-step floating catchment area (2SFCA) method: measuring the spatial accessibility to hierarchical healthcare facilities in Shenzhen, China. *International Journal for Equity in Health,* 19**,** 164.

Tarapata, Z. & Kasprzyk, R. An application of multicriteria weighted graph similarity method to social networks analyzing. 2009 International Conference on Advances in Social Network Analysis and Mining, 2009. IEEE, 366-368.

Taubenböck, H., Kraff, N. J. & Wurm, M. 2018. The morphology of the Arrival City - A global categorization based on literature surveys and remotely sensed data. *Applied Geography,* 92**,** 150-167.

Taylor, D., Klimm, F., Harrington, H. A., Kramár, M., Mischaikow, K., Porter, M. A. & Mucha, P. J. 2015. Topological data analysis of contagion maps for examining spreading processes on networks. *Nature Communications,* 6**,** 7723.

Tetteh, J. D., Templeton, M. R., Cavanaugh, A., Bixby, H., Owusu, G., Yidana, S. M., Moulds, S., Robinson, B., Baumgartner, J., Annim, S. K., Quartey, R., Mintah, S. E., Bawah, A. A., Arku, R. E., Ezzati, M. & Agyei-Mensah, S. 2022. Spatial heterogeneity in drinking water sources in the Greater Accra Metropolitan Area (GAMA), Ghana. *Population and Environment,* 44**,** 46-76.

Thomson, D. R., Gaughan, A. E., Stevens, F. R., Yetman, G., Elias, P. & Chen, R. 2021. Evaluating the Accuracy of Gridded Population Estimates in Slums: A Case Study in Nigeria and Kenya. *Urban Science,* 5**,** 48.

Thomson, D. R., Leasure, D. R., Bird, T., Tzavidis, N. & Tatem, A. J. 2022a. How accurate are WorldPop-Global-Unconstrained gridded population data at the cell-level?: A simulation analysis in urban Namibia. *PLOS ONE,* 17**,** e0271504.

Thomson, D. R., Stevens, F. R., Chen, R., Yetman, G., Sorichetta, A. & Gaughan, A. E. 2022b. Improving the accuracy of gridded population estimates in cities and slums to monitor SDG 11: Evidence from a simulation study in Namibia. *Land Use Policy,* 123**,** 106392.

Thomson, D. R. K., M.; Boo, G.; Hati, B.; Grippa, T.; Elsey, H.; Linard, C.; Mahabir, R.; Kyobutungi, C.; Maviti, J.; Mwaniki, D.; Ndugwa, R.; Makau, J.; Sliuzas, R.; Cheruiyot, S.; Nyambuga, K.; Mboga, N.; Kimani, N.W.; Albuquerque, J.P.; Kabaria, C 2020. Critical Commentary:

Need for an Integrated Deprived Area “Slum” Mapping System (IDeAMapS) in LMICs. *Social Sciences*.

Tirumala, V. R. R. S. D., Tiwari, P., Sawhney, A. & Pranatharthiharan, K. K. 2020. Analyzing Configurational Paths for Successful PPPs in Indian Urban Drinking Water Sector. *Journal of Infrastructure Systems,* 26**,** 04020023.

Tobler, W. R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography,* 46**,** 234-240.

Tockner, K. & Stanford, J. A. 2002. Riverine flood plains: present state and future trends. *Environmental Conservation,* 29**,** 308-330.

Toivonen, H., Mahler, S. & Zhou, F. A Framework for Path-Oriented Network Simplification. *In:* COHEN, P. R., ADAMS, N. M. & BERTHOLD, M. R., eds. Advances in Intelligent Data Analysis IX, 2010// 2010 Berlin, Heidelberg. Springer Berlin Heidelberg, 220-231.

Tomasiello, D. B., Herszenhut, D., Oliveira, J. L. A., Braga, C. K. V. & Pereira, R. H. M. 2023. A time interval metric for cumulative opportunity accessibility. *Applied Geography,* 157**,** 103007.

Torres, J. M., Duenas-Osorio, L., Li, Q. & Yazdani, A. 2017. Exploring topological effects on water distribution system performance using graph theory and statistical models. *Journal of Water Resources Planning and Management,* 143**,** 04016068.

Tsinda, A., Chenoweth, J. & Abbott, P. 2020. Strengthening interlinked marketing exchange systems to improve water and sanitation in informal settlements of Kigali, Rwanda. *International Journal of Urban Sustainable Development,* 12**,** 300-308.

Tsiotas, D. & Polyzos, S. 2018. The Complexity in the Study of Spatial Networks: an Epistemological Approach. *Networks and Spatial Economics,* 18**,** 1-32.

Tumwebaze, I. K., Sseviiri, H., Bateganya, F. H., Twesige, J., Scott, R., Kayaga, S., Kulabako, R. & Howard, G. 2023. Access to and factors influencing drinking water and sanitation service levels in informal settlements: Evidence from Kampala, Uganda. *Habitat International,* 136**,** 102829.

UBOS 2022. National Service Delivery Survey 2021 Repor.

UN-ECA 2014. *Assessing progress in Africa toward the Millennium Development Goals*.

UN-ESCAP 2017. Integrated approaches for sustainable development goals planning: the case of Goal 6 on water and sanitation.

UN-Habitat 2005. Situation analysis of informal settlements in Kisumu. Nairobi: UN-HABITAT.

UN-Habitat 2006. *The state of the world’s cities 2006/2007 : the Millennium Development Goals and urban sustainability : 30 years of shaping the Habitat Agenda*.

UN-habitat 2016. New Urban Agenda.

UN-Habitat 2018. SDG Indicator 11.1.1 Training Module: Adequate Housing and Slum Upgrading.

UN-Habitat 2020. Informal settlements’ vulnerability mapping in Kenya: The Case of Kisumu Settlements.

UN-Habitat 2021. SDGs indicator metadata.

UN-Habitat 2022. *World cities report 2022: envisaging the future of cities*, United Nations Research Institute for Social Development.

UN-Habitat 2023a. Rescuing SDG 11 for a Resilient Urban Planet: Executive Summary.

UN-Habitat 2023b. Rescuing SDG 11 for a Resilient Urban Planet: Synthesis Report.

UN-Habitat 2023c. The Sustainable Development Goals 11 Extended Report 2023.

UN-Water 2021. *Progress on change in water-use efficiency: Global status and acceleration needs for SDG indicator 6.4. 1, 2021*, Food & Agriculture Org.

UN‐Habitat 2004. The challenge of slums: global report on human settlements 2003. *Management of Environmental Quality: An International Journal,* 15**,** 337-338.

UN 2015a. The Millennium Development Goals Report 2015.

UN 2015b. Transforming Our World, the 2030 Agenda for Sustainable Development. United Nations.

UN 2019. *World Urbanization Prospects: The 2018 Revision*, United Nations.

UN 2020. *World Population Prospects 2019* United Nations.

UN 2023a. The Sustainable Development Goals 6 Extended Report 2023. United Nations.

UN 2023b. The Sustainable Development Goals Report 2023: Special Edition.

UN 2024. The Sustainable Development Goals Report 2024.

United Nations Office for the Coordination of Humanitarian Affairs (OCHA) 2023. Rwanda - Subnational Administrative Boundaries.

University of Rwanda 2018. Rwanda: National Urban Policies and City Profiles for Kigali and Huye.

Urfanisa, D., Kazama, S. & Takizawa, S. 2022. Evaluation of a Slum Upgrading Program for Improvement of Water Supply in Bandung City, Indonesia. *Water,* 14**,** 3025.

Usui, H. & Asami, Y. 2011. An evaluation of road network patterns based on the criteria for fire-fighting. *Cybergeo: European Journal of Geography*.

Uwayezu, E. & de Vries, W. T. 2020. Can In-Kind Compensation for Expropriated Real Property Promote Spatial Justice? A Case Study Analysis of Resettlement in Kigali City, Rwanda. *Sustainability,* 12**,** 3753.

Uwizeye, D., Irambeshya, A., Wiehler, S. & Niragire, F. 2022. Poverty profile and efforts to access basic household needs in an emerging city: a mixed-method study in Kigali's informal urban settlements, Rwanda. *Cities & Health,* 6**,** 98-112.

van der Meer, L., Abad, L., Gilardi, A. & Lovelace, R. 2021. Sfnetworks: Tidy geospatial networks. *URL: https://CRAN. R-project. Org/Package= sfnetworks*.

Velden, T., Yan, S. & Lagoze, C. 2017. Mapping the cognitive structure of astrophysics by infomap clustering of the citation network and topic affinity analysis. *Scientometrics,* 111**,** 1033-1051.

Velzeboer, L., Hordijk, M. & Schwartz, K. 2018. Water is life in a life without water: Power and everyday water practices in Lilongwe, Malawi. *Habitat International,* 73**,** 119-128.

Verboven, S., Branden, K. V. & Goos, P. 2007a. Sequential imputation for missing values. *Computational biology and chemistry,* 31**,** 320-327.

Verboven, S., Branden, K. V. & Goos, P. 2007b. Sequential imputation for missing values. *Computational Biology and Chemistry,* 31**,** 320-327.

Verplanke, J. & Georgiadou, Y. 2017. Wicked Water Points: The Quest for an Error Free National Water Point Database. *ISPRS International Journal of Geo-Information,* 6**,** 244.

Vinh, N. X., Epps, J. & Bailey, J. Information theoretic measures for clusterings comparison: is a correction for chance necessary? Proceedings of the 26th annual international conference on machine learning, 2009. 1073-1080.

Viratkapan, V. & Perera, R. 2006. Slum relocation projects in Bangkok: what has contributed to their success or failure? *Habitat International,* 30**,** 157-174.

von Luxburg, U. 2007. A tutorial on spectral clustering. *Statistics and Computing,* 17**,** 395-416.

Wagle, P. 2022. Water access disparity in Mumbai, India: Using spatial and structural attributes as formal conditionalities. *Journal of Urban Affairs***,** 1-14.

Wan, N., Zhan, F. B., Zou, B. & Chow, E. 2012. A relative spatial access assessment approach for analyzing potential spatial access to colorectal cancer services in Texas. *Applied Geography,* 32**,** 291-299.

Wang, C., Leitner, M. & Paulus, G. 2023. Multiscale Analysis of Spatial Accessibility to Acute Hospitals in Carinthia, Austria. *ISPRS International Journal of Geo-Information,* 12**,** 491.

Wang, C., Tang, W., Sun, B., Fang, J. & Wang, Y. Review on community detection algorithms in social networks. 2015 IEEE international conference on progress in informatics and computing (PIC), 2015. IEEE, 551-555.

Wang, F. 2012. Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Annals of the Association of American Geographers,* 102**,** 1104-1112.

Wang, F. 2018. Inverted Two-Step Floating Catchment Area Method for Measuring Facility Crowdedness. *The Professional Geographer,* 70**,** 251-260.

Wang, F., Antipova, A. & Porta, S. 2011a. Street centrality and land use intensity in Baton Rouge, Louisiana. *Journal of Transport Geography,* 19**,** 285-293.

Wang, F., Zheng, X.-z., Li, N. & Shen, X. 2019. Systemic vulnerability assessment of urban water distribution networks considering failure scenario uncertainty. *International Journal of Critical Infrastructure Protection,* 26**,** 100299.

Wang, H., Xiong, W., Wu, G. & Zhu, D. 2018. Public–private partnership in Public Administration discipline: a literature review. *Public Management Review,* 20**,** 293-316.

Wang, J. 2015. Resilience of Self-Organised and Top-Down Planned Cities—A Case Study on London and Beijing Street Networks. *PLOS ONE,* 10**,** e0141736.

Wang, J., Mo, H., Wang, F. & Jin, F. 2011b. Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. *Journal of Transport Geography,* 19**,** 712-721.

Wang, S. & Gu, K. 2023. Historico-Geographical Analysis of Spatial Differentiations. *In:* WANG, S. & GU, K. (eds.) *Spatial Justice and Planning: Reshaping Social Housing Communities in a Changing Society.* Cham: Springer International Publishing.

Watmough, G. R., Hagdorn, M., Brumhead, J., Seth, S., Delamónica, E., Haddon, C. & Smith, W. C. 2022. Using open-source data to construct 20 metre resolution maps of children's travel time to the nearest health facility. *Scientific Data,* 9**,** 217.

WBG 2024. Private participation in infrastructure (PPI): 2023 annual report.

Wéber, R., Huzsvár, T. & Hős, C. 2020. Vulnerability analysis of water distribution networks to accidental pipe burst. *Water Research,* 184**,** 116178.

Wei, Y.-C. & Cheng, C.-K. Towards efficient hierarchical designs by ratio cut partitioning. 1989 IEEE International Conference on Computer-Aided Design. Digest of Technical Papers, 1989. IEEE, 298-301.

Weibull, J. W. 1976. An axiomatic approach to the measurement of accessibility. *Regional science and urban economics,* 6**,** 357-379.

Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T. C. D., Howes, R. E., Tusting, L. S., Kang, S. Y., Cameron, E., Bisanzio, D., Battle, K. E., Bhatt, S. & Gething, P. W. 2018. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature,* 553**,** 333-336.

Werlin, H. 1999. The Slum Upgrading Myth. *Urban Studies,* 36**,** 1523-1534.

WHO 1992. The International Drinking Water Supply and Sanitation Decade : end of decade review (as at December 1990. Geneva: World Health Organization.

WHO 2017. Safely managed drinking water: thematic report on drinking water 2017.

WHO/UNICEF 2010. Progress on sanitation and drinking-water: 2010 update. *Progress on sanitation and drinking-water: 2010 update.*

WHO/UNICEF 2015. *Progress on sanitation and drinking water: 2015 update and MDG assessment*, World Health Organization.

WHO/UNICEF 2018. Core Questions on Water, Sanitation and Hygiene for Household Surveys: 2018 Update. New York.

WHO/UNICEF 2023a. Progress on household drinking water, sanitation and hygiene 2000–2022: special focus on gender. New York.

WHO/UNICEF. 2023b. *The WHO/UNICEF Joint Monitoring Programme Estimates on WASH* [Online]. Available: http://washdata.org [Accessed].

Wickramasinghe, A. & Muthukumarana, S. 2022. Assessing the impact of the density and sparsity of the network on community detection using a Gaussian mixture random partition graph generator. *International Journal of Information Technology,* 14**,** 607-618.

Wilson, R. 2010. *Introduction to Graph Theory,* London.

Wong, D. W. 2009. Modifiable Areal Unit Problem. *In:* KITCHIN, R. & THRIFT, N. (eds.) *International Encyclopedia of Human Geography.* Oxford: Elsevier.

Woo, B. & Jun, H.-J. 2020. Globalization and slums: How do economic, political, and social globalization affect slum prevalence? *Habitat International,* 98**,** 102152.

Woods, A. D., Gerasimova, D., Van Dusen, B., Nissen, J., Bainter, S., Uzdavines, A., Davis-Kean, P. E., Halvorson, M., King, K. M., Logan, J. A. R., Xu, M., Vasilev, M. R., Clay, J. M., Moreau, D., Joyal-Desmarais, K., Cruz, R. A., Brown, D. M. Y., Schmidt, K. & Elsherif, M.

M. 2024. Best practices for addressing missing data through multiple imputation. *Infant and Child Development,* 33**,** e2407.

World Bank 2009. Improving water utility services through delegated management: Lessons from the utility and small-scale providers in Kisumu, Kenya. Nairobi, Kenya.

World Bank 2023. World Urban Population 1960-2023.

WorldPop & CIESIN 2018. Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076). Population Counts 2000–2020 UN-Adjusted Unconstrained 100m.

Wright, J., Gundry, S. & Conroy, R. 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. *Tropical Medicine & International Health,* 9**,** 106-117.

Wu, Y., Liu, S. & Kapelan, Z. 2024. Addressing data limitations in leakage detection of water distribution systems: Data creation, data requirement reduction, and knowledge transfer. *Water Research,* 267**,** 122471.

Xing, L., Liu, Y., Wang, B., Wang, Y. & Liu, H. 2020. An environmental justice study on spatial access to parks for youth by using an improved 2SFCA method in Wuhan, China. *Cities,* 96**,** 102405.

Xu, Y., Yu, T. & Yang, B. 2019. Reliability assessment of distribution networks through graph theory, topology similarity and statistical analysis. *IET Generation, Transmission & Distribution,* 13**,** 37-45.

Xue, J., Jiang, N., Liang, S., Pang, Q., Yabe, T., Ukkusuri, S. V. & Ma, J. 2022. Quantifying the spatial homogeneity of urban road networks via graph neural networks. *Nature Machine Intelligence,* 4**,** 246-257.

Yamaoka, K., Kumakoshi, Y. & Yoshimura, Y. 2021. Local Betweenness Centrality Analysis of 30 European Cities. *In:* GEERTMAN, S. C. M., PETTIT, C., GOODSPEED, R. & STAFFANS, A. (eds.) *Urban Informatics and Future Cities.* Cham: Springer International Publishing.

Yang, S., Paik, K., McGrath, G. S., Urich, C., Krueger, E., Kumar, P. & Rao, P. S. C. 2017. Functional Topology of Evolving Urban Drainage Networks. *Water Resources Research,* 53**,** 8966-8979.

Yang, Z., Algesheimer, R. & Tessone, C. J. 2016. A Comparative Analysis of Community Detection Algorithms on Artificial Networks. *Scientific Reports,* 6**,** 30750.

Yazdani, A., Dueñas-Osorio, L. & Li, Q. 2013. A scoring mechanism for the rank aggregation of network robustness. *Communications in Nonlinear Science and Numerical Simulation,* 18**,** 2722-2732.

Yazdani, A. & Jeffrey, P. 2010. A complex network approach to robustness and vulnerability of spatially organized water distribution networks. *arXiv preprint arXiv:1008.1770*.

Yazdani, A. & Jeffrey, P. 2011. Complex network analysis of water distribution systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 21**,** 016111.

Yazdani, A. & Jeffrey, P. 2012a. Applying Network Theory to Quantify the Redundancy and Structural Robustness of Water Distribution Systems. *Journal of Water Resources Planning and Management,* 138**,** 153-161.

Yazdani, A. & Jeffrey, P. 2012b. Water distribution system vulnerability analysis using weighted and directed network models. *Water Resources Research,* 48.

Yazdani, A., Otoo, R. A. & Jeffrey, P. 2011. Resilience enhancing expansion strategies for water distribution systems: A network theory approach. *Environmental Modelling & Software,* 26**,** 1574-1582.

Yelling, J. 1982. LCC slum clearance policies, 1889-1907. *Transactions of the Institute of British Geographers***,** 292-303.

Yelling, J. I. M. 2000. The incidence of slum clearance in England and Wales, 1955-85. *Urban History,* 27**,** 234-254.

Yin, C., Liu, Y., Wei, X. & Chen, W. 2018. Road Centrality and Urban Landscape Patterns in Wuhan City, China. *Journal of Urban Planning and Development,* 144**,** 05018009.

Yin, X., Li, P., Feng, Z., Yang, Y., You, Z. & Xiao, C. 2021. Which Gridded Population Data Product Is Better? Evidences from Mainland Southeast Asia (MSEA). *ISPRS International Journal of Geo-Information,* 10**,** 681.

Yu, W., Lin, X., Zhang, W., Pei, J. & McCann, J. A. 2019a. SimRank*: effective and scalable pairwise similarity search based on graph topology. *The VLDB Journal,* 28**,** 401-426.

Yu, W., Wardrop, N. A., Bain, R. & Wright, J. A. 2017. Integration of population census and water point mapping data—A case study of Cambodia, Liberia and Tanzania. *International Journal of Hygiene and Environmental Health,* 220**,** 888-899.

Yu, W., Wardrop, N. A., Bain, R. E. S., Alegana, V., Graham, L. J. & Wright, J. A. 2019b. Mapping access to domestic water supplies from incomplete data in developing countries: An illustrative assessment for Kenya. *PLOS ONE,* 14**,** e0216923.

Yu, X., Wu, Y., Meng, F., Zhou, X., Liu, S., Huang, Y. & Wu, X. 2024. A review of graph and complex network theory in water distribution networks: Mathematical foundation, application and prospects. *Water Research,* 253**,** 121238.

Yule, G. U. 1925. II.—A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character,* 213**,** 21-87.

Zachary, W. W. 1977. An Information Flow Model for Conflict and Fission in Small Groups. *Journal of Anthropological Research,* 33**,** 452-473.

Zeng, F., Li, X. & Li, K. 2017. Modeling complexity in engineered infrastructure system: Water distribution network as an example. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 27**,** 023105.

Zhang, P., Ghosh, D. & Park, S. 2023. Spatial measures and methods in sustainable urban morphology: A systematic review. *Landscape and Urban Planning,* 237**,** 104776.

Zhang, Q., Wu, Z. Y., Zhao, M., Qi, J., Huang, Y. & Zhao, H. 2017. Automatic Partitioning of Water Distribution Networks Using Multiscale Community Detection and Multiobjective Optimization. *Journal of Water Resources Planning and Management,* 143**,** 04017057.

Zhang, X. & Tariq, S. 2020. Failure Mechanisms in International Water PPP Projects: A Public Sector Perspective. *Journal of Construction Engineering and Management,* 146**,** 04020055.

Zhang, Y. 2017. *Cluster analysis and network community detection with application to neuroscience.* University of Pittsburgh.

Zhao, F., Wu, J., Sun, H., Gao, Z. & Liu, R. 2016. Population-driven Urban Road Evolution Dynamic Model. *Networks and Spatial Economics,* 16**,** 997-1018.

Zhao, S., Liu, Y., Liang, S., Wang, C., Smith, K., Jia, N. & Arora, M. 2020. Effects of urban forms on energy consumption of water supply in China. *Journal of Cleaner Production,* 253**,** 119960.

Zhou, F., Malher, S. & Toivonen, H. Network Simplification with Minimal Loss of Connectivity. 2010 IEEE International Conference on Data Mining, 13-17 Dec. 2010 2010. 659-668.

Zischg, J., Klinkhamer, C., Zhan, X., Rao, P. S. C. & Sitzenfrei, R. 2019. A Century of Topological Coevolution of Complex Infrastructure Networks in an Alpine City. *Complexity,* 2019**,** 2096749.

Zuin, V., Ortolano, L., Alvarinho, M., Russel, K., Thebo, A., Muximpua, O. & Davis, J. 2011. Water supply services for Africa's urban poor: the role of resale. *Journal of Water and Health,* 9**,** 773-784.