Examining the Dynamics of Urban Form, Flow, and Accessibility Using Geo-Computational Methods: A Case Study of Delhi

Aviral Marwal

Department of Land Economy
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Wolfson College July 2023

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

It does not exceed the prescribed word limit for the relevant Degree Committee.

……………………………………………………………..

Aviral Marwal

July 2023

Publications

Portions of this work have been published as follows:

1. Chapter 2: Marwal, A., & Silva, E. (2022). Literature review of accessibility measures and models used in land use and transportation planning in last 5 years. Journal of Geographical Sciences, 32(3), 560-584. https://doi.org/10.1007/s11442-022-1961-1

2. Chapter 4: Marwal, A., & Silva, E. (2023). Analysing inequity in accessibility to services with neighbourhood location and socio-economic characteristics in Delhi. Journal of Geographical Analysis (Accepted)

3. Chapter 5: Marwal, A., & Silva, E.A. (2023). Exploring residential built-up form typologies in Delhi: a grid-based clustering approach towards sustainable urbanisation. npj Urban Sustainability, 3(1), 40. https://doi.org/10.1038/s42949-023-00112-1

4. Chapter 6: Marwal, A., & Silva, E. A. (2023). City affordability and residential location choice: A demonstration using agent based model. Habitat International, 136, 102816. https://doi.org/10.1016/j.habitatint.2023.102816

For each of the above listed chapters, the contribution of authors in the work is as follows –

1. Aviral Marwal: Study conceptualisation, Literature review, Data collection and preparation, Modelling, Data analysis, Paper draft writing

2. Elisabete A. Silva: Study conceptualisation, Review of draft

Left Blank

Abstract

Examining the Dynamics of Urban Form, Flow, and Accessibility using Geo-Computational Methods: A Case Study of Delhi.

Aviral Marwal

The adoration for cities is widespread across the globe. However, as urbanization escalates in the cities of the global south, concerns regarding unsustainable living have become increasingly prominent. Consequently, there is a pressing need to delve deeper into comprehending the essence of cities and their mechanisms. While understanding cities in terms of their physical configurations and the patterns of human spatial interaction has been a subject of multidisciplinary research over the past few centuries, significant advancements in the field of urban science have emerged in the last three decades. Complexity science and geo-computational models have enabled the study of cities as dynamic entities using a bottom-up approach.

This thesis constructs a conceptual framework encompassing urban form, flow, and human behaviour, which is then applied to the city of Delhi to investigate critical urban phenomena. Specifically, it examines commuting behaviour, the spatial distribution of services, typologies of built-up forms, residential location choice, and built-up expansion. In this endeavour, the study aims to provide insights into pivotal questions within urban science. These include understanding why individuals travel longer distances to their workplaces and the factors that influence their choice of travel mode. Additionally, it investigates the spatial distribution of various services throughout the city for different socio-economic neighbourhoods. The impact of urbanization on unsustainable built-up forms is also explored, along with the relationship between density patterns, and city affordability. Moreover, the study explores how urban planning can be made more efficient by incorporating the decision-making processes of planners into simulation models.

To undertake this research, diverse and novel datasets, including primary and secondary sources, were utilized for the city of Delhi. These encompassed field survey data on commuting behaviour; a spatial database containing population, income, and caste information for all residential locations in Delhi; street map data; and land satellite imageries. The study also employed various machine learning methods and spatial-statistical techniques, such as geographically weighted regression, k-means clustering, SHAP method, agent-based model, and neural network model.

The empirical findings presented in the different chapters of this thesis demonstrate that in Delhi, both urban form and flow are interconnected and influenced by human behaviour. The spatial location of households and neighbourhoods within the city plays a significant role, as does the socioeconomic makeup of these areas, in determining commuting behaviour and the spatial distribution of services. From an urban planning perspective, the city exhibits spatial heterogeneity in neighbourhood design, with the majority of neighbourhoods characterized by unsustainable built-up forms. Consequently, monitoring future built-up expansion should be a priority for Delhi's planners. Using an agent-based and neural network model, this study constructs a prioritised growth model that has the potential to showcase how planning interventions can influence future spatial growth and built-up expansion within the city. Based on the findings of this study, we recommend that future planning interventions in Delhi consider the enhancement of accessibility for low-income groups alongside environmental sustainability.

Acknowledgements

I would like to thank my PhD supervisor Professor Elisabete A. Silva for showing trust in my capabilities right from the day I reached out to her for my PhD application. Prof. Silva always encouraged me to think innovative and big and inspired me with her academic journey to excel in academic research. Her positive outlook and encouragement to my work remains instrumental in shaping the thesis to the level I could not even dream of before I began my PhD. Her feedback and guidance on the thesis and manuscripts helped me immensely to structure and articulate my work more robustly and coherently. It is due to her credit that I got excited about the field of computational social sciences and used novel research methods to study different urban issues. I will always be grateful to her for giving me the freedom to plan and execute my PhD research work the way I wanted.

I am grateful to the Department of Land Economy for admitting me to the PhD program. The flexibility of working within and outside Cambridge, especially during the COVID-19 pandemic helped me a lot to complete my PhD within my planned timeframe. The generous financial support that the department provided for my fieldwork is highly appreciated.

I would also like to express my sincere thanks to my PhD advisor, Professor Shailaja Fennell and to my viva examiners for providing me with constructive feedback on my thesis during the different stages of PhD evaluation. I enjoyed the formal/informal interactions in the department with different faculty members which also helped me to build and express my thoughts about my research work to a wider audience. I am indebted to many anonymous reviewers who played a key role in improving the quality of my research work and journal publications.

I am grateful to Wolfson College for admitting me and providing me with a homely environment throughout my PhD journey. The interaction with students from diverse research areas and social-cultural backgrounds in the college was a memorable learning experience. Different social activities in the college provided a much-sought relief and an enjoyable break from the hectic PhD schedule. I am also thankful to the college for providing me with generous conference travel funds and hardship funds.

Apart from the department and University teachings, My PhD learning has a substantial share of learning from open-source platforms on the internet from whom I learned a lot about novel research ideas and computational techniques. I have a huge respect for those who make valuable research-related content and share it publicly so that students like me can learn.

Many others have supported me in my PhD journey to whom I would like to express my sincere thanks. Professor Nomesh Bolia (IIT Delhi) has always been the source of inspiration right from my B.Tech days at IIT Delhi. I owe a lot to him for his mentoring and guidance. It is to his credit that I could develop a risk-taking attitude and the courage to pursue my dreams. I am also grateful for the invaluable life teachings that I received from Professor Anand Bulusu (IIT Roorkee), and Professor Parag Singla (IIT Delhi) during different phases of my academic life.

As I spent most of my PhD time brainstorming different ideas and their execution, the teachings from 'Bhagavad Gita' helped me a lot to stay focused and do the *nishkam karma*. I feel blessed that I was introduced to the teachings of Swami Vivekananda early in my student life, which helped me immensely to realise my potential and use it for the betterment of society. I also very keenly followed teachings from Sadhguru ji which helped me to develop an attitude of devotion toward my work and see life as endless possibilities.

The PhD journey would not have been enjoyable without the active presence of my friends. I owe a special thanks to my friends Smriti and Animesh for their support and sharing the joy, laughter, and deliciously cooked Indian food! I cherish the time I spend with my friends within the department – Joseph, Ana, Ica, Yibin, and Emily and the discussions we had while having lunch in different colleges! I am grateful for the love and support from all my friends back in India, who were always there to listen to my good and bad PhD experiences. Their presence made me cheerful, and optimistic towards my PhD.

Lastly, this PhD would not have been possible without the love and blessings of my mother, father, and sister. They stood with me as a rock in every bad phase and celebrated all the small things during the last three years. More than to my hard work and patience, I owe this PhD to them for everything they could do under the sun to make my PhD super easy and for instilling in me all the virtues since my childhood that helped me to navigate the PhD journey joyfully.

माँ और भारत माता को समर्पित...

Dedicated to my Mother and my Motherland. . . .

Table of Contents

Left Blank

List of Figures

List of Tables

Left Blank

Chapter 1

Introduction

1.1 Motivation and Study Background

The world is gradually transforming into cities. By the end of this century, a significant shift awaits us as the majority transitions from a predominantly non-urban population, which characterized the world two centuries ago, to a global scenario where urban living becomes ubiquitous (Figure 1.1). As this happens, the existing cities will get denser, new cities will emerge and the existing rural settlements will transform into urban ones.

This transformation is widely welcomed, as cities are perceived as growth engines for national economies and powerful disseminators of knowledge (NITI Aayog, 2022; Collier et al., 2018). Evidence from different countries shows that urbanisation and economic growth are correlated to a very good extent, as shown in Figure 1.2. Previous studies show that those who live in cities have higher per capita income and enjoy a better quality of life than their fellow rural citizens (Jiang et al., 2022; Yuan et al., 2020; Wang et al., 2020). Economic inequality acts as a major push factor for those living in rural areas to migrate to cities (Tiwari et al., 2022; Sulemana et al., 2019).

Beyond economic gains, cities provide improved access to resources such as quality healthcare, education, and recreational facilities, thereby enhancing human development and quality of life (Zhang et al., 2022; Tripathi, 2021). Notably, several metropolitan cities, including Mumbai, Dhaka, Mexico City, Accra, and Jakarta, exhibit higher human development indices than their national averages. Thus, cities are envisaged today as a synonym for development especially in the global south which is still predominantly rural and is characterized by a very high pace of urbanisation (UN DESA, 2018).

While future cities may enhance a nation’s economic prosperity, some concerns need to be addressed as the wave of urbanisation advances. Some of these concerns of urbanisation, as we see in the thesis's different chapters, can be related to residential segregation and ghettoization, socio-economic inequity in access to resources, wasteful commuting, negative impact on mental and physical health due to long commuting, environmental pollution caused to vehicular emissions, illegal settlements and rise of slums, and unplanned neighbourhoods. The urbanisation experience in different countries tells that economic growth in cities may not lead to inclusive growth, as it may advance growth for some at the expense of others (Kuddus et al., 2020; Dano et al., 2020). Who lives where in a city and why, and how do they commute to their workplace, are fundamental aspects of a city life that can determine whether a city is a success or failure, a magical development or tragic paralysis.

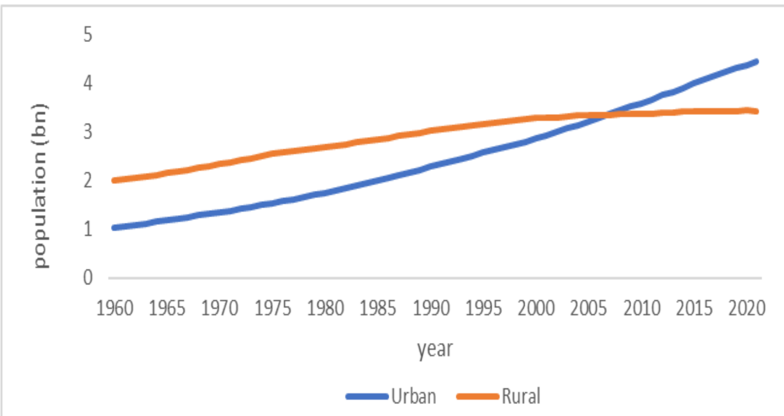

Figure 1.1: Variation in the global urban and rural population. Source: Figure created by the author, data sourced from world bank staff estimates based on the United Nations Population Division's World Urbanization Prospects: 2018 Revision.

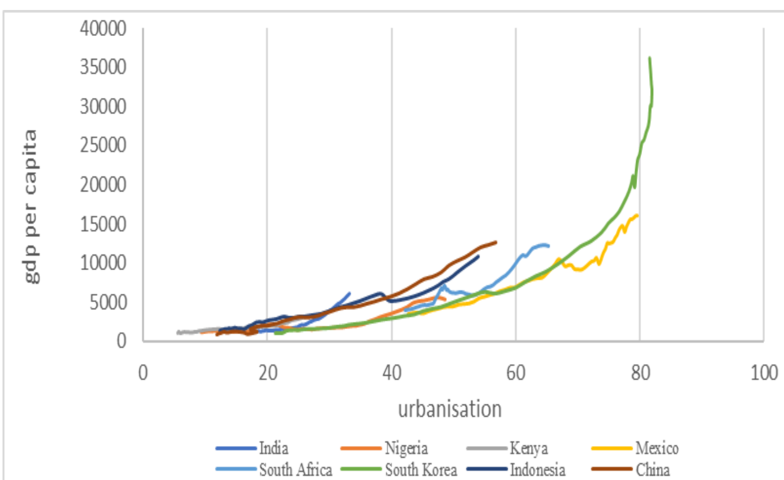

Figure 1.2: Correlation between GDP per capita and urbanisation in different countries for the years 1950 to 2016. Source: Figure created by the author, data sourced from urworldindata.org/urbanisation

As cities around the world devise strategies to cope with the ill effects of urbanization, the rise of gigapolis and cosmopolitan cities will open Pandora's box of complex questions that we have not yet considered. How we manage future cities will depend on our understanding of the processes that make up a city, or as Jane Jacobs (1961) put it, "*the kind of a problem a city is*." The experience of governing cities in the last half a century across the globe shares one common learning, i.e., cities are not mechanical structures that can be governed by a centralised planning system. Master plans and land use regulations which are implemented in a top-down manner, often fail to meet current needs and can even create more serious problems than they were designed to solve, such as the rise in slums and illegal settlements (Goytia, Heikkila, & Pasquini, 2023). With the rise of the system perspective in urban science, cities are now seen as organic entities that grow from the bottom up, incorporating the behaviour and networked interactions of thousands or millions of residents (Batty, 2013). This approach enables us to comprehend the fundamental processes governing the dynamic interactions among cities' various subsystems, thereby facilitating solutions to the "wicked problems" cities face.

1.2 Understanding Cities: A Theoretical Background

Scientific comprehension of cities and their functioning is an ongoing endeavour. Throughout the past centuries, diverse academic disciplines have examined cities using different conceptual frameworks. Notably, economic geography has proposed spatial agglomeration models, which perceive cities as physical constructs in space with a central hub for work and a transport network that delineates their boundaries. Early models, such as the Von Thunen (1826) model of central place theory, elucidated how economic activities congregate in space, resulting in urbanization and heightened profitability for consumers and producers. As economic geography progressed, more refined models of cities emerged, surpassing spatial agglomeration to describe the internal structure of cities.

Classical models of city growth, like the Burgess (1928) ring model of Chicago and the Alfred Marshall (1890) model of spatial economies, explored aspects such as residential segregation and the advantages of agglomeration. However, as cities expand, negative factors like congestion, environmental challenges, and high living costs can counterbalance the benefits of agglomeration. Henderson's (1974) concept of utility and population size equilibrium suggests that cities should strive for an optimal population size that strikes a balance between the advantages and disadvantages of agglomeration based on the dominant economic activity.

Alonso's (1964) model of land rents in a monocentric city, building upon the work of Von Thunen and Burgess, has exerted a significant influence on urban economics. It expounds on how land rents decline, and transportation costs increase as one moves away from the city centre. Other models, such as Mills (1967), Muth (1971), Wheaton (1976), and Brueckner (1987), expanded upon Alonso's work to investigate various facets of city growth and structure, integrating utility functions to model residential location preferences within budget constraints.

While classical models in urban economics could account for some important phenomena related to urban growth and built-up expansion, they were limited in their ability to incorporate the human-environment interaction as a catalyst for the dynamic evolution of cities. An important development in understanding cities occurred with the emergence of the human ecology approach (Park, 1936). According to this perspective, cities are not simply about isolated places or individuals, but rather they are ecological entities that arise from the interactions between people and the surrounding environment. Patrick Geddes considered the father of modern urban planning, remarked that a city is not merely a static physical location but rather a dynamic and ever-evolving entity with its unique narrative. To comprehend the evolution and growth patterns of cities, it is necessary to collectively analyse the different subsystems of urban environments, such as the social, economic, and environmental subsystems, and their mutual interactions.

As the human factor gained prominence in the understanding of cities, considerations of equity and justice in urban planning became important areas of study. Notable contributions in this regard include David Harvey’s (1973) "Social Justice and the City," Henri Lefebvre's (1968) "The Right to the City" (Le Droit à la ville), and Edward Soja's (2009) conceptualization of spatial justice. One central theme in these theories is that cities should plan and allocate services in a way that benefits the most impoverished and marginalized sections of society. City spaces should be designed such that they can be accessed by all, especially by those from poor socio-economic backgrounds.

Cosmopolitan cities like Mumbai, Shanghai, and New York are renowned for their diverse populations, encompassing people from various social groups. These cities offer a wide range of services that cater to the socio-economic needs of households. The existence of diverse residential settlements, including slums and illegal colonies, along with the utilization of different modes of transportation, highlights the heterogeneity in neighbourhood design, spatial distribution of services, and travel infrastructure. This symbiotic relationship between physical spaces and human behaviour imbues urban development with a dynamic character, driven from the bottom-up.

To comprehend cities from a dynamic and bottom-up perspective, a robust framework has emerged within complexity theory. According to this theory, cities can be defined as self-organizing organic systems that evolve over time. Jane Jacob (1961) described cities as "problems in organized complexity where multiple variables vary simultaneously and in a subtly interconnected way." As city sub-systems undergo constant change, the nature of their interactions also transforms, leading to a state of disequilibrium within the urban system. Traditional models depict an equilibrium state as a long-term steady state, but complexity theory recognizes the need to capture the dynamic nature of interactions among different city sub-systems and the state of disequilibrium (Batty, 2013). Simulation approaches offer a valuable tool for understanding cities in this dynamic state.

In conclusion, our understanding of cities and their functioning has advanced through various models and approaches, ranging from spatial agglomeration models to human ecology, social justice, and complexity theory. Researchers strive to capture the multifaceted nature of cities and their ever-evolving characteristics. This ongoing pursuit, employing multiple approaches, contributes to the development of a new science of cities (Bettencourt, 2021).

1.3 Conceptual Framework

This research employs a comprehensive conceptual framework to analyse cities as complex systems, offering valuable insights into their various components and their intricate interplay. The framework, depicted in Figure 1.3, provides a structured understanding of cities in terms of form, flow, and human behaviour.

Form encompasses the physical attributes of urban spaces across different scales, such as wards, neighbourhoods, streets, and households. At the ward level, it encompasses land use patterns, including residential, commercial, agricultural, wasteland, and ecological areas. Neighbourhoods-level form entails the spatial distribution of key amenities and services, such as schools, hospitals, transit stations, and job centres. At the street and household levels, form pertains to street design and building characteristics.

The second component, flow, examines commuting patterns within the city. Similar to form, flow can be analysed across different scales. At the ward level, it encompasses network and route design. The neighbourhood level focuses on transit stations and the quality of travel infrastructure, while the street level considers walking pathways and traffic congestion.

The third component, human behaviour, encapsulates individual preferences in terms of residential location choice and travel attitudes. Various factors influence these choices, including economic considerations, social and cultural norms, and environmental and health-related concerns.

According to the conceptual framework, human behaviour influences both form and flow, while form and flow also reciprocally shape each other. This interaction gives rise to diverse residential spaces and commuting behaviours within a city, which can be characterized by density patterns, socio-economic attributes, and planned or unplanned urban design. Commuting behaviour is typically examined in terms of mode choice, commuting distance, and commuting time.

Studying form and flow is crucial for assessing the quality of life provided by cities. The interaction between form and flow is manifested through accessibility. Where individuals reside and how they travel directly impact their accessibility to workplaces and various services. Achieving sustainable urban living requires urban spaces and their interactions to be environmentally sustainable and promote socio-economic equity. Thus, the framework highlights accessibility and sustainable living as vital planning objectives guiding a city's built-up expansion. Utilizing this conceptual framework, the study explores key themes and perspectives, as detailed in the subsequent sections.

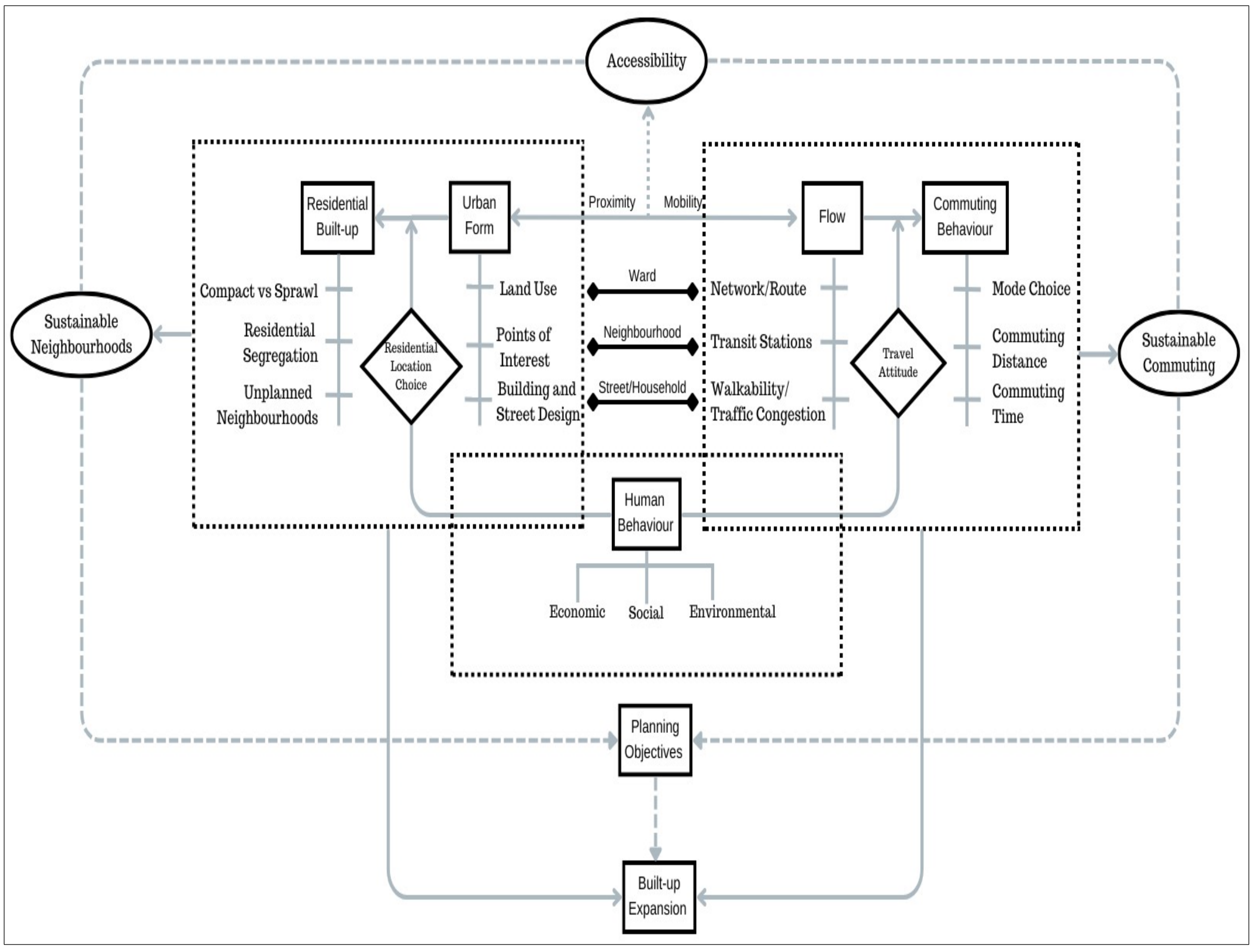

Figure 1.3: Study conceptual framework

1.4 Study Key Themes and Perspectives

The empirical work of this thesis has been carried out for the city of Delhi, India. We examine the city's form, functions, and future evolution using three vital components of cities: (1) accessibility to services, (2) built environment and travel behaviour, and (3) urban form and built-up expansion. While examining these components, we focus on issues of sustainability and socio-economic equity. Below, we introduce these components and perspectives, and the manner in which they have been examined in this thesis.

1. Accessibility to services: Accessibility is a crucial aspect of transportation planning, referring to the ease of reaching destinations. Higher accessibility enables people to reach their desired locations more quickly and effortlessly. This thesis emphasizes the benefits of high accessibility to services in a city, including improved quality of life and environmental sustainability.

Enhancing accessibility is a key concern in urban planning. It can be examined from two interconnected perspectives: mobility and proximity. Mobility relates to the ease of movement and how far one can travel within a given time frame, while proximity focuses on the nearness of destinations. Traditional transportation planning has primarily emphasized improving mobility, such as constructing new highways, bridges, and high-speed public transportation, to enable individuals to cover longer distances in less time.

However, solely focusing on mobility may not necessarily lead to improved accessibility. In some cases, people may choose to commute longer distances to reduce housing costs, resulting in decreased proximity to services. Studies have shown that despite the implementation of new travel infrastructure, commuting distances have not decreased, but instead increased in many cities (US Census Bureau, 2018; UK Labor Force Survey, 2018). Thus, while transportation planning has predominantly focused on enhancing mobility, less attention has been given to improving accessibility.

To address this challenge, transportation planning is now integrated with land use planning, which aims to improve the mix of land uses within neighbourhoods. By increasing the land use mix, people can access services in their local vicinity, reducing the need for long-distance commuting. Consequently, the key to enhancing accessibility lies in minimizing the necessity for travel and ensuring that travel, when required, is as swift as possible. In an ideal scenario, services should be brought closer to all neighbourhoods. In this context, the study analyses the spatial distribution of services to identify regions within the study area that have lower proximity to essential amenities.

As urban planning shifts its focus from prioritizing mobility to prioritizing accessibility, it is imperative to consider equity. Considering the resource constraints and welfare objectives of the government, questions regarding who benefits from increased accessibility will arise as new services are planned. Therefore, it is crucial to examine accessibility inequities as cities grow and accommodate households from diverse socioeconomic backgrounds.

2. Built environment and travel behaviour: Built environment and travel behaviour are interconnected aspects that significantly influence the functionality of a city. People engage in travel for both economic and non-economic purposes, and their travel behaviour is shaped by various factors. Extensive research has explored the relationship between travel behaviour and the built environment, using factors of travel mode choice, commuting distance, and commuting time. Empirical evidence from diverse cities suggests that the built environment plays a crucial role in determining travel behaviour.

Consequently, altering the built environment can lead to changes in people's commuting patterns, encouraging a shift towards non-motorized and public transportation. Prioritizing environmentally friendly transportation is vital for sustainable development, and cities have a pivotal role to play in achieving the ambitious goal of zero carbon emissions.

Although a correlation exists between the built environment and travel behaviour, the causal mechanisms underlying this relationship require further understanding. Cities are complex organic systems in which forms and functions emerge from the bottom-up and are influenced by individual decisions regarding residential location and travel preferences. The choice of where to live and how to commute is influenced by both the city's built environment and subjective preferences unrelated to it. These subjective preferences, known as travel-related residential self-selection (TRSS), contribute to shaping travel behaviour.

Consequently, understanding the link between the built environment and travel behaviour is more nuanced than initially perceived. Moving forward, it is crucial for cities, as intricate systems, to account for the heterogeneity in individual residential location choices and travel attitudes in order to comprehend the causal mechanisms connecting the built environment and travel behaviour. In this study, we examine the relationship between the built environment and travel behaviour among a sample of individuals in Delhi. Furthermore, we propose policy measures that can be implemented to create a city-built environment conducive to sustainable commuting.

3. Urban form and built-up expansion: Urban form encompasses the physical layout of a city, which can be observed at various scales and through different characteristics. Street-level elements such as width, curvature, slope, and footpath availability contribute to urban form, while neighbourhood or block-level factors like block area, street intersection nodes, built-up density, and land use diversity shape the overall urban form. Although previous studies have primarily focused on city-wide urban form, it is crucial to

recognize the heterogeneity within cities, especially in the global south with its diverse settlement patterns and inconsistent planning. The presence of urban slums and unauthorized colonies in metropolitan areas underscores the need to study urban form at the neighbourhood level.

Understanding urban form is essential for urban sustainability. A well-designed urban form promotes active transportation, enhances neighbourhood vitality, reduces traffic congestion, and provides sufficient housing and green spaces, all contributing to sustainable living. As urbanization accelerates in the global south, micro-variations in urban form become more apparent. Analysing spatial heterogeneity in urban form and identifying characteristics that foster sustainability is essential for effective local planning. Moreover, it expands the understanding of urbanization beyond economic and demographic growth, shedding light on whether cities in the global south unintentionally undermine sustainable living amidst rapid urbanization. This study examines how neighbourhood-level urban form features in Delhi contribute to sustainable living.

Built-up expansion is another aspect tied to urban form, reflecting the historical and future growth patterns of a city's built-up area. As cities grow, non-built-up land, such as agricultural or wasteland, is transformed into built-up areas for various purposes. However, built-up expansion is not uniform across cities. Some areas experience rapid development while others remain underdeveloped. Anticipating future growth locations and identifying influencing factors allows planners to guide growth in line with specific goals, such as promoting greener or less dense neighbourhoods.

Advancements in geo-computational models have facilitated the development of simulation techniques like cellular automata and agent-based models, enabling the monitoring and prediction of future land use expansion. While these simulations often rely on historical growth trends, it is important to acknowledge the role of top-down planning interventions in shaping urban development. Strategically planned interventions can stimulate and guide growth, and considering their impact on future development policies when simulating built-up areas provides a more accurate depiction of spatial growth. This study simulates built-up expansion in a region of West Delhi, considering future planning policies and mapping the resulting spatial shifts. By examining urban form and built-up expansion, this study aims to enhance our understanding of cities and inform sustainable urban planning practices.

1.5 Study Aim, Objectives, and Research Questions

The study primarily aims to examine the dynamics of urban form, flow and accessibility for the city of Delhi using geo-computational methods. Using spatial and non-spatial data, and quantitative methodological approaches, we examine how the different physical and non-physical components of the city influence each other, and what planning measures can be taken to make the existing city form and future urbanisation environmentally sustainable and beneficial for the different socio-economic groups. The study also aims to simulate built-up expansion in the West Delhi region using an agent based neural network model.

Objective and Questions

O 1. To study the relationship between commuting behaviour and the built environment in Delhi using findings from a household survey.

> RQ 1.1 How does the relationship between commuting distance and the built environment vary for different commuting modes?
>
> RQ 1.2 What factors are likely to bring about a change in commuters’ mode choice?
>
> RQ 1.3 Under what causal mechanism travel attitude is linked with commuting behaviour?

O 2. To examine the inequity in accessibility to services for different socio-economic neighbourhoods in the city of Delhi using a geographically weighted regression model

> RQ 2.1 How can we map the spatial variation in accessibility for different neighbourhoods in Delhi?
>
> RQ 2.2 What explains the inequity in accessibility across neighbourhoods – their spatial location or socio-economic status?

O 3 To explore the different residential built-up form typologies in Delhi using a grid-based k-means clustering algorithm and evaluate their impact on sustainable urbanisation.

> RQ 3.1 How to map and analyse the variation in the urban form at the neighbourhood level?
>
> RQ 3.2 How much residential area in Delhi can be categorised under sustainable built-up?

O 4. To demonstrate the impact of residential location choice on the urban form using an agent based model driven simulation of a hypothetical monocentric city.

RQ 4.1 How to simulate the trade-off between housing rent and commuting expenditure as reported in traditional location choice models.

RQ 4.2 Does varying rent and commuting expenditure bring spatial variation in residential density pattern?

RQ 4.3 Does the increase in income inequality and variation in land ownership intensify income-based residential segregation?

O 5. To simulate built-up expansion in west Delhi using a neural network coupled agent based prioritised growth model.

RQ 5.1 What are the urban growth driver variables associated with the built-up expansion in the studied region?

RQ 5.2 How do we utilise machine learning-based techniques to simulate non-linear relationships between urban growth and its driver variables?

RQ 5.2 How does the spatial growth pattern change and how much does the model accuracy enhance when the futuristic planning interventions are made part of the simulation model?

1.6 Outline of Research Data and Methodology

The research data and methodology have been carefully developed to align with the research objectives outlined in the previous section. The data collection process primarily involved quantitative and spatial data. To gather information on commuting behaviour and the built environment, a field survey was conducted in Delhi, targeting 1,680 households. The survey captured various aspects such as travel characteristics, socio-economic factors, household attributes, travel attitudes, residential preferences, and neighbourhood design. Chapter 3 provides comprehensive details on the survey design, while Appendix A contains the questionnaire used.

In terms of service distribution data, relevant information for Delhi was obtained from various government departments, available publicly online. This data was geocoded for spatial analysis using GIS software. Other datasets pertaining to urban form and built-up expansion were prepared by processing satellite imagery, open street map databases, and analysing statistical data from Delhi's master plans, municipal corporation, and Indian census documents. To define neighbourhood boundaries, a grid-based approach was employed using Google Earth and GIS software to map residential areas in Delhi.

The research employs a diverse range of methodologies, broadly categorized into statistical methods, land cover classification, and simulation-based methods. Statistical techniques include linear and logistic regression to investigate the relationship between commuting behaviour and the built environment. Accessibility to services is analysed through grid-based neighbourhood maps and the two-step floating catchment area method. Geographically weighted regression models are employed to study the variation in accessibility based on neighbourhood characteristics. The k-means clustering technique is utilized to cluster neighbourhoods according to dominant urban form features, providing insights into variations in accessibility, built-up density, and street design. Furthermore, the SHAP machine learning model aids in interpreting the cluster results.

The thesis incorporates two chapters that employ simulation techniques to model built-up expansion. Agent-based and neural network models are utilized for this purpose. One chapter presents an agent-based model developed in NetLogo software, demonstrating how variations in housing and transportation costs influence residential location choices and overall urban growth patterns. In another chapter, built-up expansion is simulated using three sub-models: land use classification, a neural network model to assess transition potential between land use states, and an agent-based framework to model the interaction between micro and macro agents, ultimately determining the conversion probability of cells.

Various software tools were employed throughout the study for data preparation and model execution. ArcGIS and QGIS facilitated spatial data analysis and land cover analysis. Stata was utilized for regression modelling, Python for data clustering, R for spatial regression modelling, NetLogo for agent-based modelling, and Terrset for neural network modelling.

1.7 Outline of Case Study

For this research, the city of Delhi is chosen as a case study based on a combination of objective and subjective considerations. Figure 1.4 shows the geographical location of Delhi and its ward boundaries. Delhi, the capital of India, is located in northern India, with coordinates of 28.61°N and 77.23°E. It spans an area of 1483 sq. km, measuring 52 km north to south and 49 km east to west. Administratively, Delhi is divided into 11 districts, 250 wards, and over 2,000 colonies (MCD, 2022). Over the past 40 years, the city's built-up area has expanded by more than 300%, and its population has grown from 6 million in 1981 to 11 million in 2011 (Census of India, 2011). Currently, Delhi is home to over 20 million residents and is projected to become the world's most populous city by 2030 (UN DESA, 2018). As an economic and cultural hub, Delhi attracts job seekers from across India, leading to rapid urbanization.

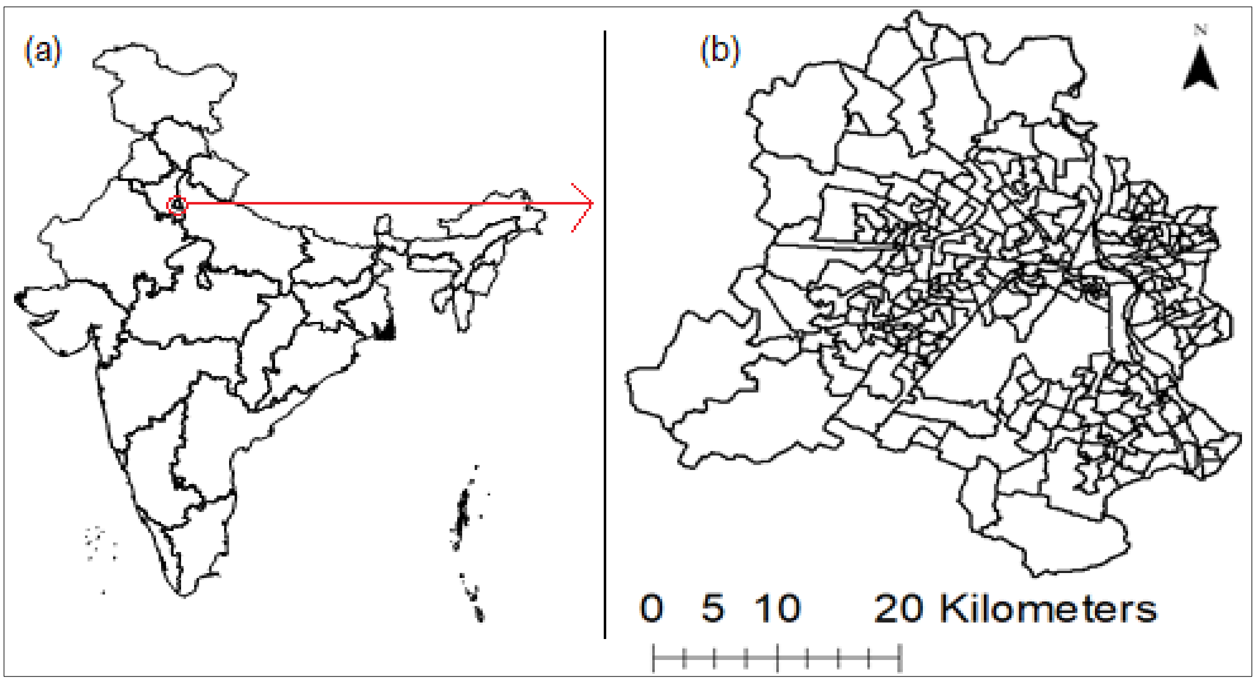

Figure 1.4: (a) Map of India with state boundaries and encircled Delhi in red, (b) Map of Delhi with ward boundaries

In the last two decades, Delhi's travel infrastructure has undergone a rapid shift. With the coming of the metro rail network in 2003, ridership in public transportation has increased, which has also impacted the residential location choice of migrants (Rana et al., 2022). Along with public transportation, the city has also witnessed a massive growth in registered private vehicles (Delhi Economic Survey, 2022-23). With an increase in built-up areas and travel infrastructure, both form and functions in the city have been changing.

In the last two-three decades, new residential areas in the city have come up to settle the massive influx of migrants from outside Delhi. These newer, low-built-up density areas have also attracted households living in Delhi's high-density neighbourhoods to move to these low-density neighbourhoods. Also, due to housing shortage and unaffordability, and inadequate planning, many of these neighbourhoods have emerged as unauthorised colonies or slums (DDA, 2022). These neighbourhoods differ in terms of their built environment, spatial location, and socio-economic characteristics. Many of these neighbourhoods lack access to basic services which is essential for a good quality of life and have poor built-up form that poses a threat to the city's environmental sustainability and individual well-being. Thus, we can say that with increasing urbanisation in Delhi, a variety of issues have emerged, such as those related to urban sustainability, neighbourhood segregation, and inequitable access to services for different socio-economic groups. On this account, Delhi becomes an important case study to examine.

While Delhi has grown immensely in the last few decades, the growth rate is expected to remain high in the coming three to four decades, given the increasing per capita income, growth in job opportunities, and preference to live in the city. This provides an opportunity to plan and channel the growth in built-up areas to create an accessible and sustainable urban form. However, given the complexity of urban form, housing, and travel, predicting future growth in built-up areas will require the use of sophisticated modelling techniques. From this perspective as well, Delhi is an interesting case to analyse.

The concerns related to sustainable urbanisation and future growth are now being examined in the literature, but few studies have explored these issues in the context of cities in the global south. Delhi, which is the capital of the most populous countries in the world today, has not been comprehensively examined considering its different aspects, such as urban form, commuting behaviour, built-up expansion, etc. The choice of Delhi as a case study adds to the growing body of knowledge on cities from the global south and provides new insights into the understanding of cities as a complex system.

1.8 Study Innovation and Contribution

The study brings innovation and significant contributions by comprehensively examining various interrelated and dynamic components of Delhi. It utilizes innovative tools, research frameworks, and datasets. The following innovative components are highlighted here and elaborated in different chapters:

Firstly, the study collects primary data on travel behaviour through a field survey and incorporates travel attitude in the research methodological framework to understand the relationship between the built environment and commuting behaviour.

Secondly, to investigate inequity in accessibility, the study creates a spatial database of 4145 residential locations in Delhi and clusters them into 1 km grid-based neighbourhoods. Socio-economic status indicators are assigned to each neighbourhood. The scarcity of spatial maps and neighbourhood-level data has been a significant challenge for studies focusing on spatial analysis in cities of the global south. This analysis uncovers inequities in accessibility at the neighbourhood level in Delhi, a previously unexplored area.

Thirdly, in the field of urban morphology, the study introduces a clustering-based methodological tool to study neighbourhood morphology. This tool allows for the dynamic and adaptable delineation of different built-up forms in the city, moving away from reliance on administratively defined boundaries. This approach provides greater flexibility in mapping and accommodates changes in the city's physical layout and demographics.

Fourthly, the study contributes significantly by theoretically modelling the bottom-up evolution of cities. It explores how individuals select residential locations to minimize travel and rent costs using an agent-based model. The simulation of different density patterns in a monocentric city under various scenarios emphasizes the importance of housing and transportation costs as spatial policy tools in shaping urban growth.

Lastly, the study's most significant and innovative aspect is the development of the prioritized growth model (PGM), which combines a neural network model with an agent-based model (ANN-ABM). Unlike previous studies, the PGM incorporates the impact of future development policies on land use simulation accurately. This approach provides valuable insights for planning agencies to shape future growth patterns.

1.9 Research Design and Thesis Structure

This section provides an overview of the chapters in the thesis and the topics they cover. The order of the chapters and their key discussions are summarized below. After the introduction chapter, Chapter 2 presents a literature review on measures and models of accessibility used in land use and transportation planning. It also examines recent literature on accessibility as a parameter for analysing travel behaviour and residential location choices. This review provides important insights and research inputs for the subsequent chapters.

Chapter 3 focuses on the relationship between commuting behaviour and the built environment for individuals working in Delhi. It introduces the case study, survey design, data collection process, and sampling techniques. The chapter then develops regression models to analyse commuting behaviour and the built environment, accounting for travel attitude and socio-economic characteristics. Policy measures for promoting sustainable commuting are discussed.

In Chapter 4, accessibility to services is measured using a potential measure, considering the cumulative count of services and supply-to-demand ratio in neighbourhoods. The analysis is conducted at the neighbourhood level, using a spatial database of residential locations in Delhi. The chapter examines how accessibility to different services varies based on socio-economic characteristics and spatial location, employing a spatial regression model. Findings inform recommendations for spatial allocation of services based on neighbourhood socio-economic characteristics.

Chapter 5 builds upon the socio-economic characterization of neighbourhoods and focuses on their physical features or urban form. Clustering algorithms are used to group neighbourhoods in Delhi and characterize them based on dominant urban forms and typologies. The chapter investigates how these diverse built-up form typologies impact sustainable urbanization, offering a fresh perspective on studying urbanization.

Moving to simulation approaches, the subsequent two chapters explore variations in urban form and built-up expansion. Chapter 6 presents an economic rational agent-based model for a hypothetical monocentric city. The model simulates the urban pattern that emerges from households' residential location choices, aiming to minimize rent and commute costs under different scenarios. It showcases the potential of agent-based models as simulation tools, revealing how urban form varies from compact to sprawl based on city affordability and the emergence of residential segregation.

Chapter 7 builds upon the agent-based model theory and applies it to a specific region in west Delhi. Historical growth patterns in the built-up area are considered, and a neural network model is used to simulate future growth in the region. The chapter incorporates the interaction between private developers and planning agencies, influencing the conversion probability of non-built-up to built-up cells. This simulation provides insights into built-up expansion in the region for 2041, considering the impact of future planning interventions.

The final chapter revisits the thesis objectives and questions, connecting them with the major findings from each chapter. It concludes by highlighting the policy implications, acknowledging study limitations, and suggesting potential areas for future research.

Chapter 2

Literature Review of Accessibility Measures and Models used in Land Use and Transportation Planning in last 5 years

Published as: Marwal, A., & Silva, E. (2022). Literature review of accessibility measures and models used in land use and transportation planning in last 5 years. Journal of Geographical Sciences, 32(3), 560-584. https://doi.org/10.1007/s11442-022-1961-1

Chapter Overview: Since its inception accessibility has undergone various changes in the way it is defined, measured, and modelled. The paper reviews the recent advancements made in accessibility measures along with the models used in different applications of accessibility related to land use and transportation. The measures of accessibility are grouped under infrastructure-based, location-based, and person-based measures. The paper finds that although the person-based measures are statistically robust and theoretically sound, they are less preferred than the location-based measure in the accessibility measurement. The review finds recent developments such as web-based mapping and the use of location-based data, image mapping through convolutional neural networks, and activity-time modelling in the measures of accessibility. Further, the paper reviews literature from the last five years that have used accessibility to study travel mode choices and household location choices and finds the use of three types of modelling framework - Statistical, Neural Network, and Agent Based models. Based on the literature review, this paper suggests the inclusion of environmental sustainability and gender equity in the accessibility measurement framework and a shift towards model synthesis to enhance the model accuracy and reduce the present complexities in model building.

2.1 Introduction

Liveable cities are key to sustainable development (UN, 2014). A liveable city is one where the local communities live a healthy life and are socially and economically prosperous. Accessibility can be labelled as the central component of liveable cities, as it impacts different activities/components of city life such as people's travel behaviour (Wolday, 2023; Van der Vlugt, 2022), residential location choice (Yan, 2020; Morales et al., 2019), social equity (Allen and Farber, 2019; Özkazanç and Özdemir, 2017), neighbourhood vibrancy (Farahani et al., 2022; Lu et al., 2019), urban growth (Van Heerden et al., 2022; Kasraian et al., 2017; Deng and Srinivasan, 2016), and environmental sustainability (Lowe et al., 2022; Lee, 2020). With cities transforming across the world, understanding accessibility for sustainable urban development has become more vital than ever before.

The use of accessibility in urban planning can be traced back to the 1920s when it was used in the location theory and transport network planning in a monocentric city pattern (Batty, 2009). From there onwards, the use of accessibility in the last 50 years has diversified to more complex city design patterns, incorporating new components and measures in the study of accessibility. With the advancement of geo-spatial techniques and the availability of microdata, accessibility is now being used to plan individual-level activity-travel patterns.

Accessibility, being a multi-faceted concept, different fields such as spatial economics, urban geography, transport engineering, architecture, urban planning, etc., have contributed to its conceptualization. Although this has enhanced the scope of accessibility, developing a unified theory and measure of accessibility has become a challenging task for researchers and planners (Handy, 2020). Accessibility has remained poorly defined and measured, which has impacted the correct usage of accessibility in its different applications (Geurs and Wee, 2004).

Several studies in the past have reviewed the developments in accessibility, the notable ones being Pirie (1979), Jones (1981), Handy and Niemeier (1997), Kwan (1998), Geurs and Wee (2004), and Paez et al. (2012). Many of the earlier reviews of accessibility have looked at the theoretical developments in the way accessibility has been conceptualized and measured (Vale et al., 2014; Wang, 2012), while some of the reviews have discussed the impact of accessibility on travel mode choices and residential location choices (Liu et al., 2020; Delbosc and Currie, 2018; Stokenberga, 2014).

The need for this literature review paper arises due to 3 reasons. First, with the advancement in geospatial technology and computation methods, the earlier measures of accessibility have been modified and thus, require a revisit. Second, new perspectives such as environmental sustainability and socio-economic equity have emerged in urban planning which can be added to the accessibility conceptual framework to make it more inclusive and comprehensive. Third, many studies in the last few years have examined the impact of accessibility on the city's land use and transportation under different models however, a review of such models is lacking in the literature on accessibility studies.

To fill this gap, our paper contributes to the existing literature on accessibility in the following two ways – First, the paper gives a conceptual context of the evolution of accessibility components and then links them with the accessibility measures. Looking at the strengths and limitations of these measures, the paper highlights how these measures can be advanced using recent technological developments such as geospatial technology and machine learning. Also, the paper suggests how the existing measures of accessibility can be made more inclusive by including the components of environmental sustainability and gender equity.

Second, the paper provides insight into how the different modelling approaches have examined the impact of accessibility on individuals' travel behaviour and residential location choices. The paper further suggests how model synthesis can overcome the limitations of these individual models and can incorporate complex measures of accessibility in a precise and meaningful manner.

The inclusion/exclusion criteria followed to select relevant literature were as follows. First, as the paper reviews only the latest developments in the accessibility measures and models, we have excluded research papers prior to the year 2015, with a maximum of our reviewed papers published during the years 2017-2020. Second, only papers written in the English Language were chosen for the review. Third, papers that have used a quantitative methodology to model the relationship between accessibility and land use and transportation, were chosen for the review.

Fourth, to limit the scope of the paper, the paper focuses on studies that have chosen transit stations to examine the impact of accessibility on travel behaviour and residential location choices. Studies which have examined the impact of accessibility to other points of interest such as schools, parks, hospitals, job destinations, etc., on land use and transportation planning, are also highlighted. The scope of the journal is kept wide to include papers from multi-disciplinary areas.

The rest of the paper is organized as follows. Section 2.2 provides the conceptual context behind developments in accessibility. Section 2.3 talks about the accessibility measures and some recent developments in it. Section 2.4 reviews the models used in accessibility applications and assesses their strengths and limitations. Section 2.5 provides the future research direction and section 2.6 concludes the paper.

2.2 Conceptual Context

This section traces the evolution of accessibility by reviewing the different perspectives and dimensions attached to it. This may appear similar to what has been done by previous studies such as Ingram (1971), Geurs and Wee (2004), and Paez et al. (2012), but our paper differs from them in the manner it has articulated the developments in accessibility in a progressive way. Accessibility is a combination of two words, access, and ability which means *"the fact of being able to be reached or obtained easily"* (Cambridge Dictionary). Hansen (1959) in his seminal paper '*How Accessibility Shapes Land Use'* defined accessibility as the "potential of opportunities for interaction", built over the concept of population potential developed by Stewart (1948). Here accessibility was seen in terms of the geographical distribution of activities, where more distant places were seen as less accessible. This gave a land use component to accessibility.

The definition of accessibility was modified as new components and dimensions were added. Ingram (1971) defined accessibility as the "inherent characteristic (or advantage) of a place with respect to overcoming some form of spatially operating source of friction". Building on this definition, Dalvi and Martin (1976) categorized sources of friction in terms of an individual's ability and behaviour, spatial variation of opportunities, and quality of transportation system. According to Dalvi and Martin (1976), the ability of the transportation system in terms of providing low-cost and high-speed travel is an important determinant of accessibility. Burns and Golob (1976) defined accessibility as the "ease with which any land-use activity can be reached from a location using a particular transport system". Thus, transportation along with land use became the two main components of accessibility, and accessibility was defined as an output of the inter-mix of the geographical distribution of activities and transportation infrastructure (Paez et al., 2012).

Both the land use and transportation infrastructure provided the spatial or geographical dimension to accessibility. Accessibility also has an aspatial or social dimension related to individual socio-economic conditions such as age, gender, ethnicity, etc. (Khan, 1992). Hagerstrand's (1975) paper "Space, time and human conditions" added a temporal perspective to the study of accessibility. In his concepts of time-geography, Hagerstrand discussed various constraints which consume an individual's time that could be allotted for different activities (Pred, 1977). This restricts the ability of the individual to reach the activity location at a specific time and diminishes his/her accessibility to that activity. Summarizing the above dimensions and components, one may say that accessibility is limited by spatial or locational constraints (distance, cost), aspatial or social constraints (age, gender), and temporal constraints (lack of time).

Taking this constraint perspective Weibul (1980) defined accessibility as an “aspect of the freedom of action of individuals”. Along similar lines, Burns (1979) proposed accessibility as the “freedom of individuals to decide whether or not to participate in different activities”. Freedom thus provides choice to an individual to choose the alternative which maximizes his/her utility. This utility-maximizing approach has given a behavioural dimension to accessibility, whereby an individual chooses that activity, location, and travel mode from which he/she can derive maximum benefit. Niemeier (1997) defined accessibility as “a value-weighted approach whose values are subjective and based on the value of opportunities assigned by individuals”. Thus, taking into account individual preferences and other barriers which restrict their movement, observed or realized accessibility may differ from the potential accessibility (Khan and Bhardwaj, 1994).

While the observed or realized accessibility occurs when there is actual utilization of services, potential accessibility refers to accessibility in the absence of any constraints (Wang et.al., 2020). Summarizing the different perspectives of accessibility, we find that accessibility has a positive correlation with the location of an activity along with individual freedom and willingness to perform that activity. Building on these observations, we propose accessibility as a degree of freedom and ability an individual has to perform a desirable activity to derive the maximum benefit from it. We interpret the word freedom as the absence of any socio-economic restrictions and the word ability as the presence of physical (bodily) and economic resources that stops or facilitate an individual to step out of his/her place and visit a location through a desired mode of travel.

Studies in past have used different components to study and measure accessibility. For example, Geurs and Wee (2004), have categorized accessibility under 4 components, i.e., Land Use, Transportation, Temporal and Individual. Building on the above discussion, this paper categorizes accessibility into 3 components - Land Use, Transportation, and Individual. All these components share either one or a mix of spatial, temporal, and behavioural dimensions. Table 2.1 highlights the accessibility dimensions under the three components. Among these three components, the individual component is specifically important as it makes accessibility a behavioural phenomenon and adds an element of individual heterogeneity in the measure of accessibility. We discuss the individual components in detail in the next sections.

Table 2.1: Accessibility components and dimensions

Component\Dimension	**Spatial**	**Temporal**	**Behavioural**
Land Use	+	-	-
Transportation	-	+	-
Individual	+	+	+

‘+’ means present,’ –‘ means absent

2.3 Review of Accessibility Measures

In this section, based on the commonly used accessibility measurement criterion which has developed in the last 50-60 years, we have divided the accessibility measures into 3 categories. Various studies such as Pirie (1979), Handy and Niemeier (1997), and Geurs and Wee (2004), have reviewed different measures of accessibility. Our review of accessibility includes measures that have not been reviewed before. The key aspect of these measures may appear to be repetitive if compared to the past reviews but our focus in this section is to show how the different measures have evolved from one another. Moreover, this section provides a review of the latest developments in the measures of accessibility, which to the best of our knowledge has not been done before in recent years.

2.3.1 Infrastructure based measures

Infrastructure based measures calculate the ease of travel by measuring the impedance occurring due to street infrastructure and availability of transportation. Street infrastructure is measured in terms of street design such as width, length, circularity, etc. The presence of cycling lanes, pedestrian lanes, footpaths, and street greenery also counts in street infrastructure measurement. A direct measure of street infrastructure is via the use of geometric calculations using graph-based or space syntax methods. The space syntax method, developed by Hillier and Hanson (1984), is a method to measure urban morphology. A topological measure of accessibility, the space syntax method uses urban morphological relationships like street integration and connectivity to measure the accessibility of a street segment to all other street segments (Huang et al., 2020). Many studies in the recent past such as Badhan (2019), Huang et (2020), Öztürk (2018), Alkamali et al. (2017), Lee et al. (2020) and Soltani et al. (2022) have used the space syntax method to characterize street networks and accessibility. A more in-depth discussion about the space syntax can be found in Yamu et al. (2020).

With the advancement of remote sensing technology and machine learning algorithms, tools such as Convolutional Neural Networks are being used to measure street design, using open street imagery platforms such as Google street view imagery. Recent studies like Weld et al. (2019), Zhang et al. (2019), Abbott et al. (2018), Choi et al. (2022), and Han et al. (2023) have used machine learning-based image classification tools to measure street quality and accessibility. An indirect measure of measuring street infrastructure is by examining the travel time, average speed, route length, road congestion levels, etc. This method is generally coupled with the location-based measures, discussed in the next section.

Another dimension related to travel infrastructure is the availability of public transportation. This may involve vehicle ownership, frequency of buses or metro rails along a particular route, waiting time at transit stations, parking time, etc. Since travel is an essential medium to reach a destination, swift travel reduces travel time with comfort to the travellers. This is a commonly used measure in transportation planning as it is easy to define and operationalize. However, this measure has one major limitation. Infrastructure based measures calculate accessibility without taking into account the need for travel. Since travel is a derived demand, measures of accessibility cannot be confined to measuring travel impedance. Features of destination location need to be added to the measure of accessibility, as seen in location based accessibility measures discussed below.

2.3.2 Location Based Measures

Location Based Measures focus on the locational attributes of the destination. They largely involve 4 types of measures based on– distance, cumulative opportunities, gravity model, and supply-to-demand ratios. Distance based measures measure accessibility in terms of proximity between two points or a group of points using linear or shortest distance between the points. Studies have also calculated the proximity in terms of non-linear distance functions such as reciprocal function, negative exponential function, and gaussian function (Ingram, 1971). For accurate measurement of travelled distance, network or path-based distance has also been used in place of Euclidean distance. Further, some studies (Vickerman, 1974; Taaffe and Gauthier, 1973) have used topological measure which calculates the number of nodes in a network in place of absolute distance between two vertices as a measure of the proximity between two points.

The cumulative measure of accessibility is an aggregate measure that counts the number of opportunities that can be reached in a specified time or that fall within a specified radius (Wachs and Kumagal 1973; Wickstrom 1971; Pirie, 1979). Based on its simplicity and easy interpretation, this is a commonly used measure of accessibility. However, it is not a precise measure as it is subject to the choice of distance/time threshold beyond which accessibility becomes zero. The choice of this threshold may vary with the region and is dependent on the choice of travel mode. Also, this measure assumes that all the opportunities are equally desirable regardless of their type (Vickerman, 1974)

Gravity based models are one of the finest measures of accessibility which combine distance based and cumulative based measures. First developed by Hansen (1959), Gravity based measures are built on Newton’s Law of Gravity, which measures potential interaction between two spatial points. They are commonly expressed as,

$$A_i = O_j \sum_{j=1} f(c_{ij}) \qquad (2.1)$$

Where A_i denotes accessibility at a point i to all other points, O_j is a factor or location attractiveness expressed in terms of size of opportunity (e.g., number of jobs) or product features (e.g., cost), $f(c_{ij})$ is the impedance function from i to j related to the travel attributes such as cost, time, or distance. Modifications in the gravity model have been done using different functional forms of impedance function such as inverse power, gaussian (Ingram, 1971), s-shaped, bell-shaped, and logistic function (Vale and Pereira, 2016; Halas et al., 2014). Among these functions, the negative exponential function (Wilson, 1971) remains the most preferred functional form.

Morris et al. (1979) pointed out that one important limitation of location-based measures is their incapability to include a factor of competition for the available opportunities in the accessibility measurement. The factor of competition comes in when the demand to access a particular opportunity is greater than what can be supplied. Thus, competition limits the number of opportunities and the opportunity attractiveness which reduces accessibility. To address this limitation, studies have included the competition factor through different indicators such as demand potential (Weibul, 1976), quotient of opportunities (Shen, 1998), and balancing factor (Wilson, 1970). More discussion about the competition factor can be found in Geurs and Wee (2004).

A widely used extension of the gravity model is the floating catchment area method (FCA). FCA measures accessibility for a location (e.g., Census tract) as the ratio of service providers to population falling in the catchment area of the location. The FCA method uses a dynamic technique of buffering in GIS to construct floating boundaries against using any fixed or administrative boundary (Peng, 1997). As per Luo and Wang (2003), one major limitation of the FCA method lies in its assumption that service in a catchment area is fully available to the locations (or residents) within that catchment area. This assumption is not always true as the services lying within a catchment area may get distributed to the locations (or residents) of other catchment areas altering the demand-supply ratio and making the potential accessibility differ from the observed accessibility.

To overcome this limitation, Radke and Mu (2000) proposed a spatial decomposition method which was simplified by Luo and Wang (2003) under the method named two-step FCA method or 2SFCA. The 2SFCA method in the first step calculates the service-to-population ratio R_j for every service location 'j' by creating a catchment area of distance threshold 'd' centred at service location 'j', as shown in Equation (2.2). In the second step, a catchment area with the same threshold distance 'd' is created for every population location 'i' and accessibility to the location 'i' is calculated as the summation of the service-to-population ratio for all the services located in the catchment area of the location 'i', as shown in Equation (2.3)

$$R_j = \frac{S_j}{\sum_{i\epsilon \left|d_{ij}<d\right|} P_i} \tag{2.2}$$

$$A_i = \sum_{j \in |d_{ij} < d|} R_j \tag{2.3}$$

S_j is the number of services at location 'j' and P_i is the population at location 'i' which falls in the distance threshold ($d_{ij} < d$). 2SFCA overcomes the limitation of traditional FCA in two ways. First, as it uses the catchment area centred at a service location in step 1, it makes all the residents in the catchment area have travel distance less than the threshold distance and thus, accessibility at location 'i' counts only those locations which fall within the threshold distance. Second, the use of a catchment area centred at a population location in step 2 makes the services to be used only by the residents within this catchment area and thus, the observed accessibility does not differ from the potential accessibility (Luo and Wang, 2003). 2SFCA has been used widely but has one major limitation i.e., it is a binary construct. It assumes every resident has equal access to a service only if they are within a catchment area and zero otherwise. Unlike the gravity model, it does not account for the impedance function (Luo and Qi, 2009).

As highlighted in Tao et al. (2020) and Jamtsho et al. (2015) in the last two decades various modifications in the 2SFCA have been done by – (a) including impedance functions such as kernel density (KD2SFCA, Guagliardo, 2004) and Gaussian (Alford et al., 2008) (b) varying the population location catchment area such as nearest neighbour method (NN-2SFCA, Jamtsho et al., 2015), base-population method (V2SFCA, Luo and Whippo, 2012), dynamic catchment sizes (McGrail and Humphreys, 2014) (c) including the supply-demand side constraints such as competition effect (Modified 2FSCA, Delamater, 2013), adjusting population demand (Luo, 2014), minimizing service demand overestimation (Enhanced 2SFCA, Luo and Qi, 2009; 3SFCA, Wan et al., 2012) and (d) incorporating different travel behaviour such as trip chaining (Commuter Based 2SFCA, Fransen et al., 2015), and use of public and private transportation (Multi-modal E2SFCA, Langford et al., 2016).

The main concern with 2SFCA and other location based measures discussed above is that they use population data of macro-level areal units which gives an aggregated measure of accessibility (Bryant and Delamater, 2019). Spatial data aggregation errors make location based measures incapable to model individual heterogeneity in terms of individual choices and preferences.

Recent advancements in location-based measures of accessibility include the use of web mapping platforms like Google Maps, Open Street Maps, and location-based services using mobile positioning data, and social network data. They are the preferred method over complex network analysis as they precisely measure the location of spatial points along with the origin-destination travel time and travel route. Using an application programming interface (API), these open-access platforms provide a real-time visualization with the accuracy of geographical data as compared to the traditional GIS based tools. They also have strong spatial analysis capabilities and can incorporate different layers such as traffic factors along different routes in the calculation of travel time. The use of web-based API can be found in recent studies like Cheng et al. (2016), Feilong et al. (2017), Niu et al. (2018), García-Albertos et al. (2018), Tao et al. (2018), Zheng et al. (2019), and Zhang et al. (2022) which have used web mapping API to calculate dynamic travel time and measure spatial accessibility to different destinations.

2.3.3 Person based measures

Person based measures can be classified into two groups – utility based and constrain based measures. Utility based measures are built on the random utility theory which assumes that an individual chooses the alternative which maximizes his/her utility. Activity based models are widely used to model the utility of activity travel in terms of choices related to the activity to be performed, destination location, travel mode, travel cost, and travel route choices. Personal and household level attributes are also incorporated in the activity-based models as these aspects affect the individual's activity-travel choices. Since choice modelling is an important component of the activity based model, discrete choice models such as Multinomial Logit Model (McFadden, 1978; Ben-Akiva and Lerman, 1985), Competing Destination Model (Fotheringham, 1986; Fotheringham et al., 2001), Nested Logit Model (Bradley et al., 2010), etc. are commonly used in the activity-based models. Maximizing the utility of an activity-travel is a way to enhance accessibility. At the same time, some constraints bound an individual in choosing a set of activity-travel.

Constrain based measures incorporate the spatial and temporal constraints of individuals in the accessibility measurement framework. Hagerstand's (1970) space-time framework provides the constraints which limit the ability of individuals to participate in different activities. Spatial-Temporal constraints in the calculation of accessibility have been incorporated using the space-time prism (STP) framework. STP is an important conceptual framework to model human behaviour. As defined by Miller (1991), "The space-time prism determines the feasible set of locations for travel and activity participation in a bounded expanse of space and a limited interval of time". Lenntrop (1976) has operationalized the space-time prism using inputs such as travel time, activity location, activity time duration, and hypothesized activity schedule to simulate the number of possible activity schedules that are regarded as the measure of accessibility (Miller, 1991). Traditional STP framework relies on geographical methods to calculate accessibility assuming constant travel speed, equitable distribution of opportunities, and using Euclidean distance measurement.

To make the STP framework more realistic, it has been refined by the addition of network based approach using GIS (Burns, 1979; Miller, 1991), cognitive constraints (Kwan and Hong, 1998), and temporal constraints (Weber and Kwan, 2002; Kim and Kwan, 2003). Recent studies such as Wang et al. (2018), Lee and Miller (2019), Zhu and Diao (2020), Fu et al. (2020), and Fu et al. (2022) have combined the temporal constraints in the activity travel based accessibility measures.

Person based accessibility measures are considered more robust and theoretically sound than infrastructure and location-based measures as they can model individual heterogeneity and spatiotemporal constraints. However, the literature highlights that the use of person-based measures in accessibility measurement is limited due to two major challenges. First, in terms of model building and operationalization, it requires micro activity and travel data which is often not available, especially in developing countries. Second, utility-based model interpretations require an understanding of complex theories which poses a challenge to their widespread applicability.

To conclude this section, we find that each of the above discussed measures of accessibility has its strengths and weaknesses. With the use of geospatial techniques and micro-data availability, it appears that measures of accessibility will be simplified in the future. No matter which measure of accessibility is used, accessibility will remain a means to an end and not an end in itself. That is, accessibility should not be seen in isolation but as an integral part of a larger socio-economic environment having ramifications on the city's economic and social development. In the next section, we review the models used in different accessibility applications.

2.4 Review of Accessibility Models on Land Use and Transportation Planning

Accessibility is a crucial parameter that is used extensively in the field of land use and transportation planning. In the context of land use planning, accessibility is considered one of the five attributes of the built environment represented through the factor of Destination Accessibility. Density, Diversity, Design, and Distance to Transit are the other four attributes of the built environment. Destination accessibility, in the literature, has been measured for different types of destinations of which jobs and transit stations are common.

This section is divided into three parts. In the first part, we review the different statistical and neural network models that have examined the impact of accessibility on land use and transportation planning using individuals' travel characteristics and household location choices in the last 5 years. In the second part, we look at a micro-simulation model – agent-based models, which have emerged as a preferred simulation tool to model the complexity of human behaviour in the literature of accessibility planning. In the third part of this section, we analyse the strengths and limitations of these models.

2.4.1 Statistical and Neural Network Models

Statistical models employ regression techniques to analyse the relationship between dependent and explanatory variables. The literature review shows that statistical models are widely used in the field of land use and transportation planning to analyse the travel mode choice or residential choice of households. Apart from accessibility, other factors of the built environment and socio-economic indicators are also used as explanatory variables in the statistical models. Based on the nature of data and study objective the statistical models have varied from simple linear regression models to discrete choice and structural equation models. In these studies, accessibility measures were found to be predominantly based on distance or cumulative opportunities.

Similar to statistical regression models, Artificial Neural Network (ANN) models can be considered a class of regression models that are used to model non-linear data. Advancements in machine learning algorithms and remote sensing technology have resulted in the wide application of Artificial Neural Network (ANN) models in the field of land use and transportation planning. ANN models consist of nodes called elementary neurons aggregated into different layers which receive inputs and convert them into outputs. Neural network modelling holds great potential in diverse applications such as ecological assessment and urban growth management (Zhang et al. 2018). Similar to statistical models, the parameter of accessibility in neural network models is also confined to two major applications - housing price and travel mode forecasting. In these applications, we found ANN models are used as a discrete choice model and are considered a better alternative to the statistical regression models.

Based on our literature review of the past five years of studies, we look into two major applications of accessibility commonly used in statistical and neural network models.

2.4.1.1 Impact of accessibility on travel

Travel characteristics are generally determined by the mode of travel, travel time, and travel route. Studies, as highlighted below, show that accessibility has a profound impact on travel characteristics. Among the travel characteristics, accessibility to transit stations has been widely studied. Literature review shows that easy accessibility to transit stations makes public transportation a preferred mode of travel. At the same time, it also develops a walking/biking culture among people. People prefer to travel by metro rail or bus when the transit station is within walking distance to their home or work location, thereby, incentivizing them to walk or bicycle to reach the transit station.

Using structural equation modelling, Cheng et al. (2020) find that origin-destination (OD) transit accessibility has a significant impact on transit mode choice while OD travel distance has no significant impact. Lu et al. (2018), using a multilevel regression model, reports accessibility to transit stations has a positive impact on walking behaviour while the other factors of the built environment have negative or no effect. Wu et al. (2020) have used a spatial regression model to find that higher accessibility to subway stations from bike stations results in more use of bikes and develop bike sharing networks. Using multivariate logistic regression, Guan et al. (2019) find higher accessibility to transit stations incentivizes households to use low-carbon transport modes such as walking and cycling.

Lee et al. (2017) find that good transit accessibility at the job (destination) site makes people use public transportation while the local urban characteristics at the trip's origin were less significant in promoting the use of public transportation. Liu et al (2016) have used structural equation modelling to find that higher accessibility to transit stations makes people choose low-carbon travel modes. Pang and Zhang (2019) use hierarchical linear models to show that better transit accessibility reduces vehicle miles travelled. Mahmoudi & Zhang (2018) use a mixed-effect regression model and find that higher drive highway accessibility discourages walking. Using structural equation modelling, Chen and Akar (2017) find that access to public transportation at tour destination make people take complex tours and travel more distances.

Apart from accessibility to the transit station, studies have found accessibility to other destinations also has an impact on travel characteristics. Using multi-level logistic regression, Lu et al. (2018) find that accessibility to retail stores and urban centres makes people prefer to walk and take public transportation while density and diversity have little effect on their commuting mode choice. Jin (2019) finds higher job accessibility decreases commuting time. Nasri et al. (2020) find job accessibility via transit contributes to the bike share demand.

In the last five years, we find that few studies have used neural network models to study the impact of accessibility on travel characteristics. Yu et al. (2016) used bus accessibility to predict the bus passenger trip flow and found the model accuracy better than the non-linear regression models. Zuo et al. (2021) used a neural network model to predict individual accessibility to bus stations. Tanwanichkul et al. (2019) built a household car ownership demand model using accessibility and other socio-economic variables and found it better than discrete choice models. Mishra and Sarkar (2017) modelled commuting choice behaviour using accessibility to public transport and found the model better than the binary logistic model.

2.4.1.2 Impact of accessibility on household location choices

Accessibility is one of the important factors that affect household location choice. Easy accessibility to attractive destinations such as job centres, parks, schools, and other public amenities plays an important role in household location choice. Earlier models of urban growth found accessibility to the central business district as an important factor in household settlement patterns. The Alonso-Muth-Mils model was one of the first such models which discussed that in a monocentric urban form, population density declines as one moves away from the central business district. With the decline in density from the city centre, the land price also declines, and thus, for low-income households, areas away from the city centre become a preferred location.

Contrary to the monocentric model, the polycentric model of new urban economics (White, 1999) recognises the existence of multiple job clusters in an urban area. The existence of multiple commercial centres and industrial centres in a city provides an effective alternative to the central business district. The polycentric model hypothesizes that population density does not decline with the increase in distance to CBD but rather increases near sub-centres (Muniz et al., 2008).

Many studies have proved that apart from distance to CBD and sub-centres, variations in household location choice can be attributed to other important factors such as proximity to urban amenities. One such frequently studied urban amenity is transit stations. Concerns about sustainable urbanization have resulted in many cities adopting policies of transit-oriented development. This has pushed the growth of public transportation especially metro rails to provide better accessibility to different destinations.

Proximity to transit stations is thus valued by households, especially by those who belong to a middle or lower economic group. AlQuhtani and Anjomani (2021) used a multiple regression model to study the impact of the proximity of residential blocks to rail transit and found that it has a positive impact on the block population density. Li et al. (2018) analysed the housing price in the inner city and suburban areas of Shanghai and found that accessibility to amenities such as parks, schools, hospitals, entertainment, etc. impacts the land price in the inner-city region while in the sub-urban areas accessibility to transit stations is valued by the property buyers.

Using the discrete choice model, Yan (2020) found transit accessibility to jobs has a positive impact on residential location choice. Saghapour & Moridpour (2019) used an ordered logistic regression model and found that public transport accessibility has a significant contribution in explaining the residential location choice of households. They note that accessibility has a greater impact than other built environment factors on the household relocation choice. Using vector autoregression, Song and Kim (2015) found that an increase in subway accessibility (due to subway network expansion) resulted in population change and land rents in the measured region. Guan and Peiser (2018) performed hedonic regression and found that metro accessibility has a significant impact on housing price which discourages low-income households from living near metro stations. Morales et al. (2017) studied the impact of accessibility to multiple destinations on land values and found that accessibility to CBD has the greatest impact on the land value compared with accessibility to other destinations. Interestingly, they found accessibility to jobs has a negative impact on land value, which they hypothesized due to the negative externalities such as pollution and congestion caused by increasing accessibility.

Within the category of household location choices, we find the use of neural network models in predicting house prices. Ruo-Qi and Jun-Hong (2020) used a genetic algorithm back propagation (GA-BP) neural network model to study the impact of accessibility to rail transit on the change in house prices. Wu et al. (2018) used an artificial neural network model to study the impact of accessibility to different public facilities on housing prices. They found the ANN model has better accuracy than the hedonic linear regression and geographically weighted regression models. Hu et al. (2019) used different machine learning algorithms including the multi-layer perceptron neural network (MLP-NN) model and found accessibility to job destinations and health centres has a significant impact on the household rental price in Shenzhen, China.

Apart from housing prices, the application of accessibility can be seen in modelling built-up areas and urban forms. Using a neural network model, Al-Sayed and Penn (2017) use street accessibility to forecast the urban form in terms of street width, building height, block density, and retail land use. Using a generalized estimating equation, Kasraian et al. (2017) found both road and rail accessibility have a significant positive impact on the urban built-up area. Similarly, studies such as Kasraian et al. (2020), and Koopmans et al. (2012) found that proximity to existing population centres has a greater impact on the built-up area as compared to accessibility to transit stations.

2.4.2 Agent based models

Agent based modelling is a bottom-up approach that micro-simulates discrete agents in an interacting environment (Babakan and Alimohammadi, 2016). The use of agent-based models in the field of land use and transportation planning has grown in the last decade with the advancement in micro-simulation techniques. Agent based models are being used to study the interrelationship between different variables such as transportation and residence location choices (Babakan and Taleai, 2015), gentrification and displacement (Eckerd et al., 2019), urban sprawl, and income segregation (Guo et al., 2017), and mobility and urban development (Leao et al., 2017). In all these studies, the thrust is to stimulate behaviour and examine the evolutionary dynamics between different agents and the environment. The agents, acting as the primary unit of study, include mainly the individuals as employees or residents. In a few studies, non-movable agents such as households and buildings have also been used (Fosset et al., 2016; Marini et al., 2019). The crucial component of an agent-based model lies in the way the behaviour of agents is modelled. This is typically done by defining a set of rules which the agents follow.

The paper finds that in most of the studies, the rules are built on a utility function or algorithm built on some statistical models like the logit model. Apart from equation-based modelling, studies elsewhere have also used cognitive frameworks such as BDI (Rao and Georgeff, 1991), PCES (Schmidt, 2002), ODD+D (Müller et al., 2013), Modelling Human Behavior (Schlüter et al., 2017) to decide agents' behaviour rules. However, to the best of our knowledge, none of the papers has used such behaviour modelling to study accessibility.

Various agent-based models have been used to study how changes in land use and transportation policies have impacted household residential choices for different socio-economic groups (Tomasiello et al., 2020). However, only a few studies have used accessibility as one of the model inputs. One of the earliest agent based residential location choice models which have used accessibility as a model input is UrbanSim (Waddell, 2000). The model uses local and regional accessibility to jobs and other facilities to simulate urban growth and real estate price. Babakan and Taleai (2015) find that the development of new transport services such as highways, BRT stations, and metro stations enhances the accessibility to different services and amenities which impacts the household rent for different socio-economic households.

Zhuge et al. (2016) used an agent based Residential Location Choice – Real Estate Price (RLC-REP) model to forecast real estate prices taking accessibility and house price as key input parameters. Tomasiello et al. (2020) used an agent-based model, ACCESS, to explore the job inequalities for different socio-economic groups. The model helps in understanding the impact of different housing and transport policies on the residential location choice and job accessibility of individuals. We now shift our focus to analysing the strengths and limitations of the statistical, neural network and agent based models.

2.4.3 Models analysis

Statistical vs ABM: Based on our literature review, we find that the application of statistical models is more widespread than agent based and neural network models in accessibility-based land use and transportation models. Statistical models offer various advantages such as robust calibration and validation techniques, and easy interpretation and application of results. However, they are not very robust in modelling the individual measure of accessibility (Hunter et al., 2018). Due to this reason, there has been a growing interest in agent based models in the accessibility literature as they can model an agent’s behaviour at a local level and thus account for the heterogeneity and complexity in human behaviour (Li and Gong, 2016). These models hold the potential to incorporate complex measures of accessibility occurring due to spatial and temporal changes in land use and transport policies (Tomasiello et al., 2020). Inspite of these advantages, we find that the use of agents-based models in applications of accessibility is still very limited.

As outlined in various studies such as Heppenstall et al. (2021), Manson et al. (2020), and Schulze et al. (2017), the key challenges in ABM lie in model parameterization, formulating agents’ behaviour rules, model calibration, and validation. What should be the appropriate behavioural rule for the agents, and should all the agents be governed by the same rule, or can there be a different rule for different groups of agents? Formulating behavioural rules requires agents' daily activity data and a very acute understanding of agents’ socio-economic conditions and their surrounding environment which is often not available. In addition, most agent-based models use the same behavioural function for every agent to model the agent’s behaviour (Dahlke et al., 2020). For example, in studies such as Tomasiello et al. (2020) and Babakan and Taleai (2015) to find optimal household location choices, all agents were modelled to maximize their accessibility to public services, which may not be a preferred choice for all agents.

As noted by Macal (2016) such a representation of agents makes the simulation unrealistic and limits the agent’s adaptability to the changing environment. Another major limitation of the ABMs is their limited capacity to calibrate and validate the results (Lee et al., 2015; Zhang et al., 2020). Due to complex model design and many parameters, calibration becomes essential so that all the parameters are fitted as per the model data. Model validation is also required to check the consistency of the model result with real-world data. As the simulation of individual activities is a very detailed and complex phenomenon, it becomes a challenging task to validate the results as a large amount of information is needed which is often not available for the entire region (Huang, 2017). According to Heppenstall and Malleson (2020), validation of an agent based model remains a “dark art at worst and haphazard at best”. This puts a question mark on the model’s efficiency and decreases the credibility of the simulation result. To strengthen the ABMs, the paper finds that studies combine them with Neural Network models.

Statistical vs ANN: The application of ANN models in land use and travel choice modelling is preferred due to their better accuracy (Shukla et al., 2016) and ability to model complex nonlinear relationships in urban design (Lee et al., 2018; Feng et al., 2015). These models are known to provide a better prediction, unlike statistical models which decrease estimation error, ANN models decrease prediction error (Tanwanichkul et al., 2019). ANN models use a hidden layer that captures the complexity or non-linearity of the dataset which statistical models are unable to do. Comparing the discrete choice analysis model with ANN models, Lee et al. (2018) find that ANN models outperform the Multinomial Logit Model (MNL) with a prediction accuracy of 80% compared with 70% for MNL.

Also, ANN models do not require many data distribution assumptions like normality and Independence to Irrelevant Alternatives (IIA) unlike their statistical counterparts (Lee et al., 2018). This makes ANN models a preferred choice over statistical models for predicting unknown data. However, statistical models are still relevant and should not be replaced completely. They give better insight into how each variable affects the model outcome, unlike ANN models which appear to operate in a 'black box' posing a challenge in result interpretability (Ha et al., 2019). In statistical modelling, it is easy to eliminate the variables which do not contribute to the model fit and there is the scope of hypothesis testing between dependent and independent variables. Furthermore, it requires a huge quality dataset to train the ANN models. These challenges in ANN models have restricted their widespread use in accessibility modelling.

2.5 Future Research Directions

This section discusses the future research direction that lies in the measures of accessibility and the models that are used to study the impact of accessibility.

(a) Gender Equity and Environment Sustainability

Today accessibility is seen as a crucial parameter of urban form. However, a value-neutral approach to enhancing accessibility can be counter-productive. The question we ask here is - Is an increase in accessibility always desirable, especially when it comes at the cost of increased environmental degradation and socioeconomic inequity? At the conceptual level accessibility needs to be defined and measured by including its impact on the environment and social-economic equity. With the availability of high-resolution spatial data, the use of disaggregated measures or person based measures of accessibility is going to increase.

While many of the accessibility measures incorporate individual preferences to travel, they do not explicitly capture the inequity in accessibility arising due to socio-economic factors such as age, gender, income, etc. Ignoring such individual characteristics in accessibility measures masks the inequity that exists within different socio-economic groups (Dixit and Sivakumar, 2020). Limiting the scope of the paper to one of the socio-economic factors, we found that only a few studies have examined gender inequity in accessibility measures. A study by Lecompte and Pablo (2017) in Bogota shows that women spend more time commuting than men for the same distance and have lower job accessibility per capita. They further suggest that this accessibility inequity becomes stronger in lower socio-economic groups.

Kwan and Kotav (2015) in their survey in Bulgaria find that the daily travel time of women is higher than men because women in the sample used public transportation instead of a private vehicle as their primary mode of travel.

In our literature review, we found the measures of accessibility so far developed are largely male biased or at best gender neutral as they do not factor explicitly women's perception towards the mode of travel and destination location which can be different from men. The case of gender-based inequity in accessibility is very important to measure especially in patriarchal societies where women face different barriers to travel. To make the accessibility measures more inclusive they need to be modified taking into account those factors which impact women's preference of travel mode and travel location. This includes transport safety, security, comfort, reliability, cleanliness, and factors like a perceived threat of violence or harassment at the travelled location (Pirra et al., 2021). Thus, women’s accessibility differs from men’s accessibility as it is not just confined to factors of time, cost, and availability of opportunities but also includes other factors of safety and security to which women give more preference than men (Lecompte and Pablo, 2017).

Apart from gender equity, our review suggests environmental sustainability is a potential area of research in accessibility studies. As discussed in a previous section different perspectives on accessibility – transportation, land use, and individual - have evolved in the last 60 years. Today, with an increase in vehicular emissions and their negative impact on the environment there is a need to have an environment-centric approach to accessibility as the environmental issues have been either avoided or not addressed explicitly in past studies of accessibility (Kinigadner et al., 2021; Määttä-Juntunen et al., 2011). The concerns of environmental sustainability in the study of accessibility arise as an individual’s utility maximization approach can have a negative impact on the environment (Johansson-Stenman and Martinsson, 2006).

To maximize accessibility, an individual may choose a travel mode that provides the highest utility in terms of travel cost and travel time. However, the choice of travel mode also decides the extent of damage to the environment occurring through vehicular greenhouse gas emissions and other pollutants (Inturri et al., 2017; Woodcock et al., 2007). Thus, the environmental cost of the travel mode (or the external travel cost) should be taken into account in the measure of accessibility along with different travel attributes like time, cost, and comfort (Kinigadner et al., 2020; Vasconcelos and Farias, 2012). If environment sustainability becomes a component of accessibility measures, then non-motorized travel will contribute to enhancing accessibility scores. However, the overall accessibility with non-motorized travel might still be less if the travel time or cost involved in such travel modes is higher than the motorized travel.

(b) Model Synthesis

Looking at the scope of these models, recent studies show that model efficiency can be enhanced with the coupling of statistical / machine learning (ML) models with agent-based models. Gore et al. (2016) built a novel approach to statistical debugging to enhance the efficiency of trace validation and verification of agent-based models. Zhang and Vorobeychik (2019) in their review of different categories of agent-based models discuss the use of machine learning models to calibrate and predict human behaviour in agent based models. They find that although only a few studies have incorporated machine learning models in ABM simulations, advancement in data analytics is a positive development to solve the issue of model calibration and validation.

Carrella et al. (2019) use a simple linear regularized regression to calibrate an agent based model. Studies such as Lamperti et al. (2018), and Zhang et al. (2020) have performed calibration of ABM through surrogate modelling techniques using machine learning algorithms. By combining techniques of intelligent iterative sampling and machine learning, a surrogate of the ABM is built which makes the calibration of the model easier and less time-consuming. Crooks et al. (2020) highlight how different studies have used machine learning models at different stages of agent-based model formulation such as model parameterization to set decision rules for the agents or to carry out model optimization and estimations.

In their review of agent-based models using machine learning algorithms, Dahlke et al. (2020) find that agents' behaviour can be made more realistic by making them learn their behaviour during the simulation through the use of Multi Agent Reinforcement Learning (MARL). Some studies in the energy sector such as Kofinas et al. (2018), and Wang et al. (2019) have used a Q-learning algorithm to carry out Multi Agent Reinforcement Learning. Edali & Yucel (2018) have used random forest metamodels and sequential sampling to understand the input-output relationships in the agent-based simulation models.

Similarly, through the use of other ML algorithms such as genetic algorithm deep nets, decision trees, and inverse re-enforcement learning, studies have obtained better realistic simulations (Van der Hoog, 2019; Ramchandani et al., 2017; Negahban, 2017; Laite et al., 2016). The use of such ML algorithms makes agents learn and adapt their behaviour to changing environmental conditions. This suggests that model synthesis can bring significant changes in making the model result more efficient and robust.

We end this section with a cautious note about the use of ML in ABM. ML models will certainly enhance the explanatory power of ABM in the future; however, the focus should not be on making the model grand but on its simplification and meaningfulness. That will happen only when modellers keep an eye on the model processes and not just on the model efficiency. Understanding *what goes in*, *what comes out*, *and what happens in between* should be the guiding torchlight in ML-based simulations. To put it in other words, modellers using ABM should understand the need to use the ML and the associated cause and effect in the simulation, otherwise as noted by Dahlke et al. (2019), the use of ML can result in the creation of "intelligent yet black box ABMs".

2.6 Conclusion

As cities are complex entities that evolve organically with time, a critical understanding of accessibility measures and models related to different applications of accessibility can help policy makers to solve many challenges of urban development. This literature review was aimed at understanding how accessibility has been measured and modelled in the past. The review found that many studies have defined accessibility using different dimensions and components such as land use, transportation, spatial and temporal. The study defines accessibility as "the degree of freedom and ability an individual has to perform a desirable activity to derive the maximum benefit out of it". The study reviewed the measures of accessibility and categorised them under infrastructure based, location based, and person based measures. Within each such measure, recent advancements were highlighted.

Three notable advancement includes - the use of machine learning tools to measure street forms, the use of web-based mapping software and location-based data to measure precise travel and activity location, and accounting temporal variations in activity based measures. The measures of accessibility have a normative purpose as they specifically focus on what policy planning measures can be taken to improve accessibility. The study suggests the use of environmental sustainability and gender equity as crucial factors that should be included in the measure of accessibility, as a future research direction.

The paper reviewed two major applications of accessibility along with the models used in these applications. The models are used as an explanatory tool to study the impact of accessibility on travel mode and household location choices. Many studies report that accessibility to transit stations makes public transportation a preferred mode of travel and is considered an important factor in household location choice.

Noting the strengths and weaknesses of these models, statistical models are found to be widely used as they are statistically robust and easy to interpret but do not account for heterogeneity at the individual level. Neural Network models are preferred over statistical models as they account for non-linearity in the data and have better prediction accuracy. Challenges of big data requirements and interpretability of results restrict the wide use of neural network models in accessibility modelling.

Agent based models are known to simulate the heterogeneity in individual behaviour but have issues related to behaviour simulation, calibration, and validation. The paper finds that studies in the recent past have combined statistical/neural network models with agent-based models. This suggests that model synthesis can bring significant changes in making the model result more efficient and robust. With easy access to micro-data and advancement in modelling software, complete integration of the three models can result in better operationalization of the accessibility measures in its different applications. Such an approach appears to be growing in the accessibility literature and thus, holds potential for future research.

Chapter 3

Commuting Behaviour, Built Environment and Travel Attitude: Insights from a Field Survey in Delhi

Chapter Overview: Limited studies in the global south have analysed the residents' commuting behaviour due to a lack of publicly available data on travel behaviour. The study examines the relationship between commuting behaviour and built environment for working residents in Delhi controlling for their socio-economic and household characteristics, and travel attitude. A field survey was performed in Delhi interviewing 1679 respondents and the data was analysed using a descriptive statistical approach and regression modelling. The model results show that some built environment characteristics such as spatial location of household and workplace, have considerable influence on commuting distance and mode choice. However, socio-economic characteristics such as income and gender were found to have a stronger influence on determining commuting behaviour. Importantly, the study shows that travel attitude and residential self-selection have a role in explaining the causal linkage between built environment and commuting behaviour which has been less examined in previous studies from the global south. Based on the survey findings, the study recommends policy measures to make commuting in the city sustainable.

3.1 Introduction

Commuting plays a crucial role in individuals' economic and social lives. People commute to work for economic reasons, but their commuting behaviour reveals much about their travel preferences, socio-economic status, and the city's infrastructure. Research indicates that understanding why individuals travel long distances to work and the factors influencing their choice of transportation can significantly improve their health and well-being (Liu et al., 2022; Chatterjee et al., 2020). It can also help address emerging issues related to unsustainable urbanization, such as vehicle emissions, traffic congestion, noise pollution, and residential segregation (Ashik et al., 2023; Stein et al., 2022; Zhu et al., 2022).

While there are hundreds of studies examining commuting behaviour and its determinants, most of them have been from the cities of the global north, specifically from North America and Europe. There is a scarcity of empirical evidence on this issue from high-density cities in the global South (Acheampong, 2020; Van Wee et al., 2019). Delhi is currently experiencing one of the highest rates of urbanization (UN DESA, 2018) and population growth globally, leading to changes in the city's urban form and travel infrastructure. The expansion of the metro rail network and the significant increase in metro ridership over the past decade have had a substantial impact on the commuting behaviour of city residents (Bhat, 2022; Tayal & Mehta, 2021). As the capital of the world's second-most populous nation, Delhi has unfortunately received limited attention in terms of studying its commuting behaviour. Therefore, it becomes crucial to consider Delhi as a case study to enhance our understanding of the relationship between the built environment and commuting behaviour, and to broaden the scope of research in this field.

The study aims to analyse the impact of the built environment on commuting behaviour after controlling for socio-economic and household characteristics, and self-selection bias of commuters in Delhi. We study commuting behaviour using the two widely used indicators namely, commuting distance and mode of commuting. The study becomes significant as its findings are based on the primary data collected through a household survey of 1679 working individuals residing across the city. Apart from the built environment characteristics measured at residential locations, the study also included the travel attitude and workplace location characteristics in the analysis of commuting behaviour which have been less examined in the context of the global south (Ding & Cao, 2019).

The rest of the paper is organised as follows. Section 3.2 provides a brief literature review on the relationship between commuting behaviour and built environment factoring in travel attitude. Section 3.3 provides the study context and survey design. Section 3.4 discusses the survey findings and gives an overview of secondary data used in the study. Section 3.5 builds a descriptive analysis of the survey findings. Section 3.6 builds a regression model to examine the influence of the built environment and other factors on commuting distance and mode choice. Section 3.7 concludes the paper.

3.2 Literature Review

Given the widespread impact of intra-city travel, there is a substantial body of literature examining commuting behaviour and its relationship with various aspects of urban development, often utilizing travel survey data (An et al., 2022; Guan et al., 2020; Sun & Yin, 2020). By considering previous studies, one can categorize them based on the study context, methodological approach, and main findings.

Built environment characteristics that influence commuting behaviour can be categorized into three spatial scales: city-level (distance of households from the city centre, location of employment hubs), neighbourhood-level (population density, proximity to amenities), and street-level (street width, intersection density). While the impact of these characteristics on commuting behaviour may vary depending on the study's context and methodology, a general trend can be observed in the literature: in neighbourhoods with higher density and mixed land use, people have shorter commuting distances. At the same time, neighbourhoods with good access to public transit incentivise people to commute longer distances. Recent studies support these findings (Wang et al., 2022; Gupta et al., 2022; Zhu et al., 2022; Nakshi & Debnath, 2021; Shin, 2020).

Additionally, some studies have examined the influence of street design variables on commuting behaviour. They have found that streets with safe sidewalks and a higher density of intersections encourage active forms of commuting and use of public transport (Nabipour et al., 2022; Gaglione et al., 2021), while wider streets promote car usage (Yang et al., 2021). Other aspects of the built environment, such as the distance of households and workplaces from the city centre, have also been found to significantly impact mode choice and commuting distances (Duquet & Brunelle, 2020; Zhu et al., 2020).

In addition to the built environment, studies have shown that the socioeconomic and household characteristics of commuters also play a significant role in influencing commuting behaviour. Income is consistently identified as one of the key factors determining commuting behaviour, with higher-income individuals more likely to use cars and commute longer distances to work (Guerra et al., 2022; Nkeki & Asikhia, 2019). Commuting behaviour has also been found to vary based on gender, age, and education levels (Kersting et al., 2021; Havet et al., 2021; Guan & Wang, 2019; Cheng et al., 2019). Household characteristics such as ownership, size, and the presence of children have also been associated with commuting behaviour (Chidambaram & Scheiner, 2021).

While the impact of the built environment on commuting behaviour has been studied in relation to socioeconomic and household characteristics, researchers have increasingly recognized the importance of considering commuters' travel attitudes in establishing a causal relationship (Deng & Zhao, 2022; Van Wee & Cao, 2022). By controlling for travel attitudes, recent studies show the impact of the built environment on travel behaviour can be more accurately determined, leading to more realistic policy frameworks. However, previous studies, in the context of the global south, have limited inclusion of travel attitudes in their modelling and analysis of commuting behaviour (Zhu et al., 2023; Nakshi & Debnath, 2021).

Travel attitudes can directly influence travel behaviour, such as when commuters who prefer to minimize travel costs are more inclined to use public transportation. Additionally, travel attitudes can indirectly impact travel behaviour through the mechanism of travel-based residential self-selection (TRSS) (Van Wee et al., 2019; Wolday et al., 2019). Residential self-selection in this context refers to the tendency of commuters to choose a residential location that aligns with their travel preferences. For instance, individuals who prioritize comfortable commuting may opt to prefer using a car for their commutes, as opposed to those who aim to minimize travel expenses and select a residence close to transit stations.

Ignoring residential self-selection can lead to an overestimation of the impact of the built environment on travel behaviour (van Herick & Mokhtarian, 2020; Kroesen & Chorus, 2018; Liu et al., 2018). However, residential self-selection may also be influenced by social inequality, as households may choose residential locations based on income and housing affordability rather than their travel preferences (Zhang et al., 2019). In such cases, disregarding residential self-selection does not significantly bias the effect of the built environment on travel behaviour, particularly when socioeconomic and demographic variables are considered (Wolday, 2018).

Considering the various factors discussed above that influence commuting behaviour and the issue of causality, the study constructs a conceptual framework, depicted in Figure 3.1. The framework illustrates that when residential self-selection (RSS) is driven by travel attitudes, the chosen built environment reflects the individual's travel preferences. In this case, the built environment should not be considered as the primary cause of travel behaviour. However, when RSS is influenced by socio-economic or non-travel-related reasons, the built environment can be regarded as the primary cause of travel behaviour. Conceptually, it should be acknowledged that travel attitudes are built by various psychological beliefs and social and cultural factors.

Previously some studies taking Delhi and other Indian cities as a case study have examined the relationship between travel behaviour and the built environment. However, these studies had limited scope as only a few of them considered travel attitudes and other subjective factors related to the built environment in their research design (Patnala et al., 2023). Additionally, many of these studies did not take into account the diverse characteristics of the built environment at different spatial scales, particularly at the street level.

Given the limited consideration of travel attitudes in existing studies from the global south, this study addresses this gap by incorporating the travel attitudes of the respondents as control variables in the regression modelling. The study also explores the reason behind respondents' choice of travel mode and their residential location choices through descriptive data analysis. Due to data constraints, this study does not include factors of residential location choice in the regression model. However, by including these important variables as part of descriptive data analysis, the study aims to provide a more comprehensive understanding of the factors influencing commuting behaviour and the role of residential self-selection and travel attitudes in this context.

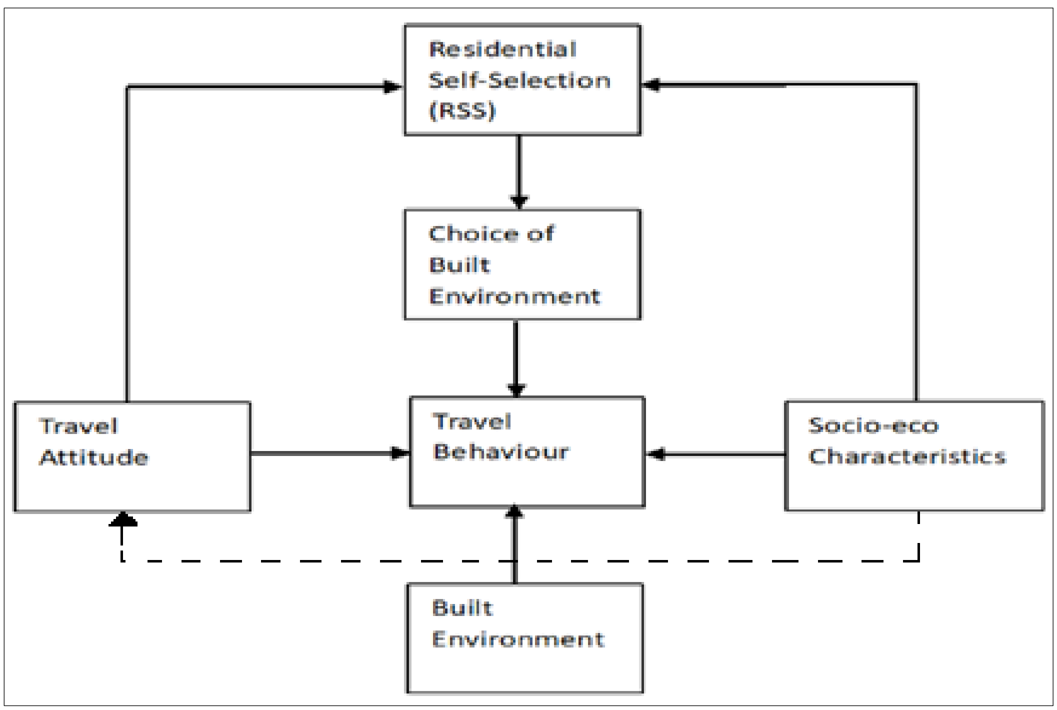

Figure 3.1: Causal mechanism between the built environment and travel behaviour

To the best of our knowledge, this study is the first to provide an in-depth analysis of commuting behaviour and the built environment in Delhi. We believe that the study's findings can provide valuable insights for planning measures aimed at promoting sustainable and public transport-oriented commuting practices.

3.3 Study Context and Survey Design

Delhi is located in northern India, with coordinates of 28.61°N and 77.23°E. It spans an area of 1483 sq. km, measuring 52 km north to south and 49 km east to west. Administratively, Delhi is divided into 11 districts, and 250 wards, and has over 2000 colonies (MCD, 2022). Over the past 40 years, the city's built-up area has expanded by more than 300%, and its population has grown from 6 million in 1981 to 11 million in 2011 (Census of India, 2011). Currently, Delhi is home to over 20 million residents and is projected to become the world's most populous city by 2030 (UN DESA, 2018).

As an economic and cultural hub, Delhi attracts job seekers from across India, leading to rapid urbanization. The introduction of the Metro rail system in 2003 has had a significant impact on the commuting pattern of the city residents' (Rana et al., 2022). The availability of an efficient public transportation system in the city resulted in an influx of workers from other parts of the country to settle in the city's outer areas, transforming the land use and physical form of Delhi and its nearby areas, collectively known as the National Capital Region (NCR) (Naikoo et al., 2020).

According to the Delhi Economic Survey 2022-23, the number of private vehicles (cars and motorcycles) in Delhi has more than doubled from 317 per thousand people in 2005-06 to 643 per thousand in 2019-20. In 2021-22, there were nearly 8 million registered private vehicles in the city, with two-wheelers accounting for around 67% and private cars for 28% of the total. Between 2015 and 2020, the number of private cars increased by 13% and two-wheelers by 35% in Delhi. Despite the rise in personal vehicle use, public transportation ridership in the city has also increased. From 2015 to 2020, metro ridership grew by 6.3% and in 2019-20, the average daily passenger ridership on the metro was 2.8 million.

Survey Design

The household survey aimed to gather information on residents' travel behaviours and their influencing factors. It included data on travel characteristics, socio-economic and household characteristics, built environment, travel attitude, and residential location preferences. The survey targeted full-time employed individuals with fixed workplaces who commuted at least four days a week. Unemployed individuals, those working from home, or those travelling outside the city for work were excluded. Each household was represented by one respondent, with priority given to female workers. Figure 3.2 shows the map of Delhi and the locations of the surveyed households

Neighbourhoods in Delhi were selected for interviews based on economic status. From a list of approximately 2,500 residential locations classified into eight categories by land price data, 200 locations were preliminarily chosen to ensure equal representation across all categories and distributed throughout the city. Within each selected location, 8-10 households were randomly chosen for interviews, with a minimum distance of 100 meters between them. Consideration was given to matching the household design with the typical design of the neighbourhood.

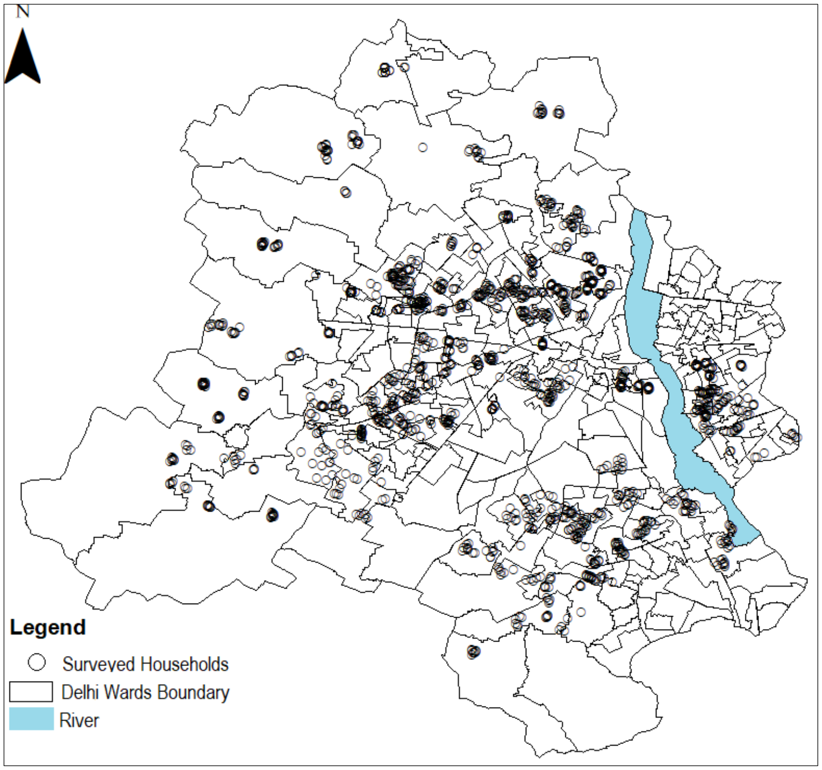

Figure 3.2: Household surveyed in Delhi

Since respondents were employed, interviews were on weekends when they were available at home. The majority of the interviews were face-to-face and few were conducted telephonically. The survey was carried out in two phases, from October to December 2021 and February to March 2022. Fifteen field surveyors were selected from Delhi based on their academic background and previous survey experience and were trained on the questionnaires and interviewee selection criteria. Before the actual survey, they conducted five dummy interviews to familiarize themselves with the questions. A pilot survey was then conducted to assess the interviewees' understanding and make necessary modifications to the questionnaire.

During the survey, ethical guidelines and data collection norms were followed. Respondents were informed about the survey's purpose and only those who voluntarily agreed to participate provided written consent. The participation form assured respondents that the data would be used solely for academic purposes and would not be shared with any other agency. The survey was conducted using the Google survey application on smartphones. Data quality was checked immediately after completion, and any issues were addressed by conducting follow-up interviews with the households. Interviews typically took 15-20 minutes, and each surveyor interviewed 8-10 households per field day. The overall response rate was close to 25%. The low response rate was mainly because many households refused to talk due to privacy concerns and fear of getting the COVID-19 virus. A total of 1679 households were surveyed, but some initial surveys were excluded due to missing information on household travel attitudes. Therefore, the final analysis included 1456 households.

The study acknowledges one of the known biases in survey data collection which is social desirability bias. It is a common form of response bias that occurs in primary data collection, particularly in surveys and interviews. It refers to the tendency of respondents to provide answers that they perceive as socially acceptable or favourable rather than responding truthfully. This bias can distort research findings and lead to inaccurate or overly positive or socially desirable responses. Keeping the bias into consideration, we tried to minimise it through the following measures: We did not ask respondents their names and assured them that their responses would remain anonymous and would not be shared with anyone else and would be used only for study purposes. Some questions were framed indirectly or using less direct language to reduce the pressure on respondents to provide socially desirable responses. Our field surveyors were trained to critically observe the household conditions and match them with the responses on household economic status such as income, number of vehicles, etc. Finally, we made sure that the respondents were comfortable answering the questions and were familiar with the purpose of our survey so that they responded genuinely.

3.4 Survey Data Findings

3.4.1 Socio-economic and household characteristics

In our survey of 1,679 people in Delhi, 24% were women, and the average age of all respondents was 39 years. We found that fewer women participated in our survey, mainly because they had concerns about their privacy and safety when sharing information about their lives and travel habits. Additionally, the low rate of women participating in the workforce in Delhi (~15%) contributed to the smaller number of female respondents available for interviews. The majority (74%) were married and 52% had children above the age of 3. Education-wise, 53% were graduates, 28% had post-graduate degrees, and 19% had completed high school or below. The average household size was 4-5 members, with 2 working members. In terms of income, categories included high income (16%), high middle income (22%), middle income (27%), lower middle income (24%), and low income (11%). Employment types were primarily private salaried jobs (53%) and self-employment (30%). Vehicle ownership-wise, 37% of households had no cars, 42% had one car, and 21% had multiple cars. With regard to the dwelling type, 61% of the respondents lived in terraced houses that run along a residential street, while the rest of the respondents lived in a flat or apartment-based residential society. Considering household ownership, 26% of the respondents lived in rented accommodations. Approximately 35% of respondents were those who have shifted their residential location from outside or within Delhi in the last 10 years and are referred to as migrants in this study.

3.4.2 Built environment characteristics

The study takes into account the built environment characteristics at the city level, neighbourhood level and street level. One of the most commonly studied built environment characteristics is the distance from the city centre. In the context of Delhi, the city centre is the location where the city of Delhi first came into existence and from where the built-up expansion gradually spread to outer areas. While the city has become polycentric with the growth of different business and commercial centres, referred to as district centres, the city centre is still a place of different kinds of employment and recreational activities with a well-developed transport infrastructure connected with the rest of the city.

In our survey, 40% of respondents were located in the city's inner area having a distance of less than 10 km from the city centre and 46% of them were located in the city's outer area with a distance from the city centre in between 10 to 20 km, the rest, 13% of respondents were located in city periphery, i.e., beyond 20km from the city centre. Additionally, the proximity of households to their nearest district centre was also measured.

Workplace location characteristics were also considered, including the distance from the city centre. Results showed that 30% of respondents worked in the city's inner area, 43% in the outer area, and 27% had their workplace more than 20 km from the city centre. The study also explored the distance from workplaces to metro stations and parking availability, as these factors can influence workers' travel behaviour (Yan et al., 2022; Islam & Saphores, 2022).

Characteristics such as population density and accessibility to transit stations were considered at the neighbourhood level. To determine the neighbourhood population density, we conducted a spatial mapping of all residential locations in Delhi, along with their respective population counts. The population data for residential locations were obtained from the Delhi Election Commission website (SEC, 2022). For each surveyed household, the neighbourhood population density was calculated by summing the population of all residential locations within a 1 km radius of the interviewed household. Table 3.1 illustrates the mean population density of interviewed households, showcasing a decreasing trend as the distance from the city centre increases.

To assess the proximity of households to metro rail stations, we initially geocoded the locations of all metro stations in Delhi. Subsequently, using the near tool in ArcGIS, we measured the household's proximity to the nearest metro station. Bus stop density was determined by summing the number of bus stops located within a 1 km radius of the surveyed households.

Table 3.1: Mean population density in the surveyed household’s neighbourhoods

Household Distance from City Centre (km)	Mean Population Density (per sq. km)	% of respondents
0 - 5	32205.26	5.54
5 - 10	25913.74	34.43
10 - 15	24290.46	31.51
15 - 20	25385.76	15.01
20 - 25	6253.75	7.5
25 - 30	4356.03	6.02

Finally, at the street level, street intersection density was studied due to its impact on traffic congestion and discouraging personal vehicle use (Ding et al., 2022). Using the Delhi residential road network from OpenStreetMap (OSM), we measured the street intersection nodes within a 1 km radius of surveyed households in the ArcGIS software. Street design aspects like width, footpath availability, and slope for the streets just outside the surveyed house were recorded during field surveys by the field surveyor. We also asked the respondents about their perception of traffic jams in their neighbourhood, and on street cleanliness and safety on a scale of 1 (worse) to 5 (best). 24% of the respondents felt that their neighbourhood has frequent traffic jams. Around 13% and 11% of respondents gave a score of 1 or 2 in street cleanliness and street safety, respectively.

3.4.3 Travel characteristics and self-selection bias

To analyse commuting patterns, respondents were asked about their commuting distance, time, and mode. The average commuting distance reported by respondents was 11 km, with an average commuting time of 34 minutes. To understand the dominant mode of travel, respondents were asked to specify their primary mode used for most days in a week, covering the longest distance. Table 3.2 presents mode choices, average commuting distance, and time. Among the options, two-wheelers were the most common mode (31%), followed by cars (29%). Public transportation options included Delhi metro rail, and public and private buses, while non-motorized modes encompassed walking, cycling, and battery-operated rickshaws.

Table 3.2: Average commuting distance and commuting time for different modes

Mode of Travel	**Number of Respondents (in %)**	**Average Commuting Distance (km)**	**Average Commuting Time (minutes)**
Car	489 (29)	12.59	34.75
Two-wheeler	517 (31)	8.41	27.27
Public Transport	400 (24)	14.92	48.34
Non-Motorized	153 (9)	1.96	13.66
Cab	100 (6)	17.1	41.85
Shared Auto	20 (1)	6.37	20.4
Total	1679		

The study acknowledges the self-selection bias in mode choice and commuting distance in terms of travel attitude and residential location choice. We parametrise travel attitude under three factors namely, cost minimising, time minimising, and comfort maximising. Although previous studies have used a multi-response approach to construct different dimensions and then extract the major dimensions as attitude-governing factors, in this study owning to survey constraints, we directly asked the respondents to rank the three factors which are considered important in defining travel attitude. These factors were ranked as, 1 (very important), 2 (somewhat important), and 3 (not so important). A higher rank to a factor showed a higher degree of importance to that factor over the other factors while selecting the commuting mode. We find that 51% of the respondents ranked commuting time as very important, while 27% ranked commuting expenditure as very important. The remaining 22% ranked commuting comfort as very important.

We also included responses on residential location choice in our survey questionnaire. Respondents were asked to rank their preference for their desired residential location proximity to seven different locations if they were given a chance to relocate. The facilities included - workplaces, public amenities (schools, hospitals, parks, etc.), transit stations, highways, commercial centres, religious places, and similar social group neighbourhoods. Proximity to these locations was ranked on a scale of, 1 (highly preferred) to 7 (least preferred) in a hierarchic order. The distribution of the highest and lowest rank and the average rank for all these locations is shown in Table 3.3. It shows that proximity to workplaces and public amenities is the most preferred residential location choice, followed by transit stations. Proximity to roads and commercial centres is less preferred while proximity to religious centres and neighbourhoods of similar social communities are the least preferred residential location choice for the households.

Table 3.3: Preference to different residential locations and the average rank given to these locations by respondents.

Residential location choice in proximity to:	Percentage of respondents		Average rank
	Rank 1 (Very highly preferred)	Rank 7 (Least preferred)	
Workplace	52.27	2.68	2.05
Amenities	30.36	1.58	2.21
Transit Staitons	5.29	8.52	3.52
Roads	1.63	9.67	4.58
Commercial Centres	4.83	3.8	4.05
Religious Places	0.9	34.28	5.84
Social Community	3.26	29.02	5.73

Table 3.4: Summary statistics of the variables.

Variable	Mean	Std. dev.	Min	Max
Socioeconomic and household characteristics				
Age	38.15	10.31	17	88
Gender (1 = Female)	0.24	0.43	0	1
Education (1 = Graduate or above)	0.82	0.39	0	1
Income (1 = Upper middle class or above)	0.39	0.49	0	1
Household Members	4.58	1.42	1	12
School going children (1 = Yes)	0.53	0.5	0	1
Migrant status (1 = Yes)	0.33	0.47	0	1
Household plot area (sq.m)	100.92	84.47	8.28	828.1
Dwelling (1 = Flat)	0.34	0.47	0	1
Ownership (1 = Rented)	0.26	0.44	0	1
Travel characteristics				
Home distance to workplace	10.86	10	0.1	55
Mode Choice (1 = Car)	0.29	0.45	0	1
Mode Choice (1= Motorbike)	0.31	0.46	0	1
Mode Choice (1 = PT)	0.24	0.43	0	1
Mode Choice (1 = NMT)	0.09	0.28	0	1
Travel frequency to workplace (1 = More than once a day)	0.12	0.33	0	1
Built environment characteristics				
Population density	22391.42	16184.8	408.92	92723.89
Home distance to metro station (km)	2.96	2.58	0.08	13.57
Bus stop density near residence	22.85	12.86	0	76

Home distance to CBD (km)	12.8	6.18	1.54	28.38
Home distance to district centre (km)	4.83	4.07	0.18	17.78
Workplace distance to CBD (km)	16.05	8.8	1	48
Workplace distance to metro station (1 = More than a km)	0.56	0.5	0	1
Parking availability at workplace (1 = Yes)	0.61	0.49	0	1
Street width (feet)	14.61	6.25	3	40
Street intersection density	484.93	241.8	35	1372
Footpath availability (1 = Yes)	0.4	0.49	0	1
Jam frequency (1 = High)	0.24	0.43	0	1
Cleanliness (1 = Poor)	0.11	0.31	0	1
Safety (1 = Low)	0.08	0.27	0	1
Travel Attitude				
Minimising travel expenditure (1 = very important; 3 = not so important)	2.13	0.8	1	3
Minimising travel time (1 = very important; 3 = not so important)	1.57	0.64	1	3
Maximising travel comfort (1 = very important; 3 = not so important)	2.29	0.79	1	3

While the residential location preferences showcase a general preference based on which a household may select their future residential location, they do not translate into their present choice of residence as these preferences may not necessarily have been valued that way by the respondents when choosing their current neighbourhood. Due to this, their residential location preferences do not correlate theoretically with their commuting behaviour. Hence in this study, we do not factor the residential location choice as an explanatory variable in predicting commuting behaviour. However, the issue of residential self-selection is examined using open-ended survey questions, as discussed in the next section.

A data summary of all the variables used to study commuting behaviour is shown in Table 3.4

3.5 Descriptive Analysis

In this section, we explore the role of residential self-selection, travel attitude and built environment on commuting behaviour by descriptively analysing the survey data.

To understand how residential self-selection (RSS) can be a factor in determining commuting behaviour we asked the survey respondents, who have changed their residential location within Delhi in the last ten years, to disclose the primary reason for doing so. Among our survey sample size of 1679 respondents, 408 reported that they had shifted their residential location within the city itself. Summarising the reasons for residential relocation, 35% of them reported –'*to have a bigger house* or *to have their own house'*, as the primary reason for relocation. 19% of them reported *proximity to the workplace as their reason to* *relocate,* while *12%* of them relocated *to lower household rent.* For 13% of the respondents, social issues such as *marriage* or *preference to live within similar social groups* were responsible for their relocation. *Proximity to transit stations* and other amenities was reported by 8% of the respondents as the reason behind relocation. The choice of *low-density neighbourhoods* and choice of *cleaner environment* was evident in 6% and 5% of the respondents, respectively.

In the above analysis, travel-related RSS is evident from the commuters' choice of living near workplaces and transit stations. This shows that people select their residential location based on their preference for commuting distance and travel mode. However, travel-related RSS is visible only in the case of a minority of respondents, and other factors especially, the preference to have a bigger house or to have one's own house are major factors behind relocation which does not directly correlate with commuting distance or travel mode choice. This shows that travel-related RSS does play a role in influencing commuting behaviour however, its influence appears to be smaller as compared to other reasons of RSS.

An important point of consideration is to understand what causes households to choose residential locations near workplaces or transit stations. A possible explanation is their travel attitude or preference to minimise commuting costs or time. We explore this linkage between travel attitude and commuting behaviour while analysing the result of the regression model.

To study the role of built environment on commuting behaviour we draw on the analysis of our survey findings that is built on the two structured questions which we asked our respondents. First, respondents who used private transport were asked about their reason for not using public transport for daily commuting to the workplace and second, respondents who have changed their mode of commuting to the workplace in the last ten years were asked to disclose their reason for doing so.

Out of the total survey sample size of 1679 respondents, 1279 respondents did not use public transport for commuting to the workplace. Considering the question of why the respondents choose the private mode as compared to the public mode for commuting to the workplace, our survey findings hint at two main reasons. First, 31% of the respondents reported commuting by metro was inconvenient because of poor connectivity from their household or workplace to the metro station. The average distance from residence to the nearest transit station for the non-public transport users and public transport users was found to be 3.2 km, and 1.8 km, respectively. This shows high accessibility to transit stations is important in incentivising the use of public transportation.

Second, as reported by 27% of the respondents, the reason for not using public transport was their household proximity to the workplace which does not require them to use the metro or bus. Thus, when commuting distance is less, commuters find the use of public transport has lower utility over private transport.

Other important reasons were as follows. 10% of the respondents felt that commuting by public transport is time-consuming and thus, preferred to travel by car or two-wheeler. 9% of the respondents could not use the metro/bus owing to the nature of their job which involves visiting clients or field-based tasks or late-night working hours, which makes commuting by metro more cumbersome. 7% of the respondents do not use public transport as they find it overcrowded and suffocating and 6% of the respondents use office vehicles for commuting to the workplace. The remaining 10% of the respondents did not prefer public transport due to different reasons like health issues, privacy issues, joint travel with spouse or colleagues, and higher travel expenditure.

The above findings show that, in the case of the majority of respondents, by enhancing the proximity to transit stations the use of public transport in daily commuting can be enhanced. This directs us towards further exploring the role of the built environment in influencing commuting behaviour.

We now examine the change in mode choice for our respondents. We asked respondents if they had changed their commuting mode to the workplace in the last ten years. Out of the total surveyed 1679 respondents, 615 reported that they have changed their commuting mode in the last ten years, out of which 365 respondents reported that they did not relocate to any other residential location and workplace within or outside Delhi. This means all these 365 respondents had a job and were commuting to their workplace in the last ten years and did not change their residential and workplace location.

Table 3.5 shows the count of respondents who have changed their mode choice, the rows represent the earlier travel mode, and the columns represent the new travel mode. The below analysis is for the sample of 365 respondents.

Table 3.5: Change in mode choice for the studied sample

Earlier/New	Car	Two-wheeler	Metro	Bus	Cab/Auto	NMT	Total (%)
Car	-	7	13	0	5	9	34 (9.31)
2-wheeler	51	-	17	6	5	13	92 (25.2)
Metro	36	21	-	5	9	4	75 (20.54)
Bus	11	49	20	-	9	8	97 (26.57)
Cab/Auto	13	9	10	2	-	0	34 (9.31)
NMT	1	27	1	1	3	-	33 (9.04)
Total (%)	112 (30.68)	113 (30.95)	61 (16.71)	14 (3.83)	31 (8.49)	34 (9.31)	365 (100)

The survey result shows that of all those who have changed their travel mode in the last ten years without changing their workplace and residential location, close to 31% of them have shifted to car from other travel modes, of which shift from two-wheelers has the maximum count followed by metro. The majority of these respondents who shifted to car for daily commuting to the workplace reported an *increase in income* as the primary reason for the change in travel mode. Some also stated that commuting by car *saves time* and for a few others, car is a better option concerning their health issues. Another equally big shift in the travel mode lies where close to 31% of the respondents have shifted to two-wheelers from different travel modes, of which shift from the bus is the highest followed by non-motorised transport, such as walk, bicycle or rickshaw. In this scenario also we find, that an *increase in income* and a *decrease in travel time*, are the reasons for the change in travel mode from bus/NMT to two-wheelers.

The third highest shift can be seen toward the use of the metro comprising close to 17% of the respondents. The majority of these respondents have shifted to the metro from bus followed by two-wheelers. When inquired about the reason for this shift, the majority of them reported *operationalisation of a new metro station* near their residence or workplace prompted them to use the metro to commute to the workplace. Few of them also found that commuting by metro takes less time compared to their previous mode of travel. We also notice that commuting by bus remains the least preferred choice as less than 4% of the respondents who changed their mode choice in the last ten years shifted to commuting by bus. On the other hand, 27% of the respondents who have changed their mode choice shifted away from commuting by bus to a new travel mode.

Summarising the three important factors for change in travel mode, an increase in income was reported as the main cause by 30% of the respondents, while close to 17% of the respondents changed their travel mode to decrease travel time. Shift in travel mode due to the availability of transit stations was observed in 9% of the respondents. Long commuting distances, travel expenditure, and health issues were also stated as the cause of a shift in travel mode. This shows that the preference to minimise travel time is an important factor in influencing the change in the mode of travel. Also, the built environment is a factor, although minor in influencing the change in mode of travel.

3.6 Regression Modelling

3.6.1 Built Environment Influence on Travel Distance

We use linear regression to model the relationship between commuting distance and built environment, as shown in Equation 1,

$$Y_i = \beta_0 + \beta_1(BE_i) + \beta_2(SEH_i) + \beta_3(TA_i) + \varepsilon \quad (3.1)$$

where Y_i refers to the commuting distance of respondent i; BE_i are the set of variables measuring built environment of location where the respondent i lives and works; SEH_i are the set of variables measuring the socioeconomic and household characteristics of respondent i; and TA_i refers to variables measuring the travel attitude of the respondent i.

The model hypothesizes that built environment indicators, except for street design, influence commuting distance after considering socio-economic and household characteristics and travel attitudes. Street design is excluded from the statistical model since it's unlikely to impact commuting distance. To account for variation by travel mode, as shown in previous studies (Zhu et al., 2020), we analyse three sub-models: one for all respondents, one for public transport users (bus and metro), and one for private transport users (car and two-wheelers). Non-motorized and cab-based commuting are excluded due to the limited sample size. Model results are presented in Table 3.6, with 3 sub-models showing their coefficients and standardised coefficient values, denoted as beta.

Among the considered seven built environment variables, three variables were found to have a significant influence on commuting distance in the combined model. As shown in sub-model (3), a one-unit increase in distance to the city centre decreases commuting distance approximately by 0.4 units. That shows that within the groups of respondents who commute by private transport, those who are located in the city's inner areas commute longer distances to the workplace as compared to those who live in the city's outer area. This means that the job housing balance is perhaps lower in the city's inner area. Evidence of this finding can be related to the emergence of IT and management consulting companies in the nearby towns of Delhi namely Gurugram and Noida in the last two decades (Dutta et al., 2020; Kushwaha & Nithiyanandam, 2019). However, the impact of household distance to the city centre on commuting distance is insignificant in the case of respondents using public transport, as shown in sub-model (2).

Another important finding is the 0.4 units increase in the commuting distance with a unit increase in the workplace distance from the city centre. This shows that respondents whose workplace is situated in the city's outer areas commute a larger distance than those whose workplace is in the city's inner area. This holds for both travel modes, however, the elasticity is higher for private transport users.

The household distance to metro stations also has a significant impact on the commuting distance in the combined model. It shows that commuting distance decreases as one lives away from the metro station. In other words, those living near the metro station commute longer distances. This is possible as living near to metro station increases the likelihood of the use of the metro which makes it convenient to cover longer distances. However, we do not find any significant relationship between them when separately analysed for the different travel modes.

Table 3.6: Linear regression model result

Dependent Variable = HH Distance to workplace	**Sub-Model (1): All Modes**		**Sub-Model (2): Mode Public Transportation**		**Sub-Model (3): Mode Private Transport**	
	Coefficient	Beta	Coefficient	Beta	Coefficient	Beta
Socioeconomic and Household Characteristics						
Income (1 = High Income)	1.41** (0.6)	0.06	3.06** (1.38)	0.12	1.85** (0.73)	0.1
Age	-0.10*** (0.02)	-0.10	-0.13** (0.05)	-0.12	-0.03 (0.03)	-0.04
Gender (1 = Female)	-0.27 (0.59)	-0.01	-1.35 (1.14)	-0.06	-1.48* (0.82)	-0.05
Education (1 = PG or above)	4.39*** (0.68)	0.1	3.80** (1.65)	0.12	2.51*** (0.84)	0.10
School going children (1 = Yes)	-2.48*** (0.51)	-0.12	-0.77 (1.17)	-0.03	-1.92*** (0.62)	-0.10
Household members	0.26 (0.18)	0.03	-0.15 (0.42)	-0.02	0.37 (0.23)	0.05
Migrant (1 = Yes)	-0.95* (0.58)	-0.04	1.003 (1.22)	0.04	-1.22** (0.71)	-0.06
Dwelling (1 = Flat)	0.98* (0.59)	0.04	0.56 (1.25)	0.02	0.70 (0.75)	0.03
Ownership (1 = Rented)	-0.92 (0.68)	-0.04	-4.35*** (1.26)	-0.20	0.411 (0.931)	0.01
Household Area	-0.001 (0.003)	-0.001	0.00 (0.009)	-0.01	0.006 (0.003)	0.06
Built Environment Characteristics						
Population Density	0.000 (0.000)	0.04	0.00 (0.0)	0.02	0.000 (0.000)	0.05
HH distance to city centre	-0.31*** (0.06)	-0.19	0.16 (0.14)	-0.102	-0.38*** (0.08)	-0.25
HH distance to district centre	0.01 (0.13)	0.006	-0.25 (0.28)	-0.09	-0.02 (0.16)	-0.01
Workplace distance to city centre	0.35*** (0.03)	0.34	0.30*** (0.07)	0.24	0.37*** (0.04)	0.35
HH distance to metro station	-0.43*** (0.16)	-0.11	0.53 (0.48)	0.08	-0.16 (0.19)	-0.05
Workplace dis. to metro station (1 = More than a km)	0.28 (0.51)	0.01	3.92 (1.08)	0.18	-0.09 (0.63)	-0.004
HH Bus stops density	-0.03 (0.02)	-0.03	0.00 (0.05)	-0.009	-0.03 (0.02)	-0.05
Travel Attitude						
Cost Minimising over Time Minimising	0.89* (0.59)	0.03	0.901 (1.17)	0.04	0.76 (0.80)	0.03
Comfort Maximising over Time Minimising	1.11* (0.63)	0.04	2.70* (1.55)	0.09	1.30* (0.72)	0.06
_cons	8.846*** (2.07)	.	14.09*** (4.19)	.	6.43** (2.66)	.
Number of observations	1456		349		876	
R-squared	0.27		0.29		0.24	
Adj R-squared	0.25		0.23		0.21	
Root MSE	8.91		9.38		9.43	

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

Model results show some of the socioeconomic and household characteristics have a significant influence on commuting distance. Respondents with high income and education levels commute longer distances which shows that respondents aspiring for high-paying and more qualified jobs are less likely to find such workplaces locally and thus, need to commute longer distances. We find females tend to commute shorter distances as compared to males, which is an expected finding and often reported in the literature. Age and the presence of school-going children in the household negatively influence commuting distance which is also a common finding in the literature. An interesting finding to note here is that respondents who have migrated to the city in the last ten years commute a lesser distance than the non-migrants. This shows that possibly migrants have the opportunity to choose their residential location closer to their workplace. These findings reverberate the reasoning that we discussed in our earlier descriptive analysis section, showcasing that residential self-selection can be a causal factor behind variation in the commuting behaviour of the respondents.

The nature of dwelling types also influences commuting distance, as we find that those living in flats or apartments tend to commute longer distances than those living in independent houses. In the context of Delhi, this is an important finding because the city has primarily independent houses, however, in the last couple of decades there has been a rise of flat-based housing societies in the city's peripheral areas where people may not have many workplaces options making them commute longer distances.

To study the influence of travel attitude we compared the attitude of cost minimisation and comfort maximisation with time minimisation. The attitude of time minimisation was chosen as the base category as the model hypothesis that travel is a derived demand, whereby people travel for a purpose and would like to minimise their travel journey and time. Also, there exists a trade-off between commuting cost and time, as one minimises commuting cost it may enhance the journey time (as slower travel modes are often cheap).

The model results show that respondents who preferred minimising the commuting cost over minimising the commuting time, commute long distance. Those who prefer to maximise commuting comfort over minimising commuting time also tend to commute longer distances. These results hint at some statistical association between travel attitude and commuting distance and suggest that there exists some underlying causal mechanism that relates these two factors. Possibly, a particular travel attitude may influence the residential location choice of households which then decide their commuting distance. For example, an individual who prioritises minimising commuting cost over commuting time is more likely to use public transportation. Given the higher utility of public transportation in commuting longer distances, such an individual may choose his/her residential location away from the workplace, resulting in an increase in commuting distance.

Similarly, those who prioritise commuting comfort are more likely to travel by car. As car commuters are usually from high-income households and for them, residence in the city’s inner area is affordable and a favourable choice due to the high accessibility to other public services in these locations. As we discussed above, such commuters living in the city’s inner area commute longer distances. An in-depth explanation of how travel attitude influences commuting distance will require analysing the time series data on residential location choice and workplace choices of households.

Based on our above analysis, we find that commuting distance is explained by different factors having different degree and nature of impact, that varies under public and private transport. While the built environment characteristics especially, workplace distance to the city centre and household distance to the city centre have the strongest influence on commuting distance, our model shows that travel attitude is also an important factor that may shape the residential location choice and thus, have an indirect influence on bringing variation in the commuting distance.

The theory of spatial agglomeration, often linked with economic geography and urban economics, extensively examines urban growth and commuting patterns in cities. Policies promoting spatial agglomeration have shown benefits, including positive externalities and increased employer profits. However, they also bring about negative externalities, such as longer commutes, higher commuting costs, and increased traffic congestion. With the rise of information and communication technology (ICT) since the 1980s, a lively debate has emerged regarding the relevance of agglomeration economies in the 21st century. On one side, proponents declared the 'death of distance' (Cairncross 1997) and the emergence of dispersed metropolises (Mitchell 1996). Conversely, others argued that ICT growth would amplify agglomeration economies (e.g., Gottman 1982; Gillespie 1992).

Our findings on commuting distance show that Delhi is a case of ‘increased agglomeration economies’. With growth in ICT in the last two decades along with the rise of metro rail infrastructure offering cheap commuting, workplaces have not dispersed as many would have expected, rather have agglomerated at two locations namely Noida and Gurugram, lying outside the city boundary. The pattern of outward commuting, as visible in our study, shows the influence of these job clusters in attracting Delhi’s work force. In this context, we can say that in Delhi agglomeration economies will continue to have an impact on the choice of workplace location and thus, urban policies should not disregard agglomeration in the era of the digital world. Rather, urban policy in Delhi should focus on creating multiple employment clusters prioritising jobs related to the IT sector that can result in the creation of local agglomeration economies which have advantages of scale to employers and give more choice to job seekers in choosing their travel mode choice and residential location.

3.6.2 Built Environment Influence on Travel Mode

To account for variation in the travel mode choices with the built environment, we employ a binomial logistic regression model that provides the likelihood of the occurrence of a particular travel mode for a given set of independent variables. The model hypothesises that the built environment characteristics including the street design elements have a significant influence in enhancing the likelihood of a travel mode after controlling for socio-economic and household characteristics and travel attitude. We model the likelihood of commuting by car, two-wheeler, and public transportation, as shown in Table 3.7.

The influence of built environment characteristics was found to be more significant on the likelihood of commuting by car and public transportation as compared to two-wheelers. We find, that an increase in workplace distance from the city centre increases the likelihood of commuting by car and decreases the likelihood of commuting by public transportation. This signifies that people rely less on public transport and more on cars for commuting to workplaces that are located in the city's outer region. This is an expected outcome as we find Delhi although has a good public transportation network within the city's inner areas but lacks coverage in the city's outer areas. This makes the commuters working in the city's outer areas rely more on cars for commuting. This requires that workplaces located in the city's outer areas need to be made more accessible to the transit stations.

Considering the proximity to metro stations, we find commuters living away from the metro station are more likely to use a car for commuting to the workplace and less likely to use public transportation. Similarly, those whose workplace is in proximity to metro stations are more likely to use public transportation. This is an important finding that relates to the importance of transit-oriented development in promoting the use of public transportation. A similar finding was reported in the descriptive analysis section where we showed that some of the respondents changed their commuting mode to the metro with the operationalisation of metro stations in their neighbourhood.

Parking availability at the workplace also enhances the likelihood of commuting by car which shows that the built environment at the workplace can also influence the mode choice. We find that the level of population density does not have any significant relationship with the commuting mode choice. This finding is in contradiction to many previous studies that show higher population density enhances the likelihood of commuting by public transport (Altieri et al., 2020).

Table 3.7: Logistic regression model result

Travel Mode	**Car**	**Two-wheeler**	**Public Transport**
Socioeconomic and Household Characteristics			
Income (1 = High Income)	1.673*** (0.183)	-0.862*** (0.174)	-1.173*** (0.202)
Age	0.036*** (0.008)	-0.02*** (0.007)	-0.017** (0.008)
Gender (1 = Female)	-0.825*** (0.204)	-0.985*** (0.187)	1.071*** (0.176)
Education (1 = PG or above)	0.884*** (0.313)	-0.294* (0.173)	0.732*** (0.237)
School going children (1 = Yes)	0.834*** (0.175)	-0.251* (0.140)	-0.253 (0.163)
Household members	-0.069 (0.067)	0.06 (0.051)	0.035 (0.057)
Migrants (1 = Yes)	0.089 (0.192)	0.021 (0.156)	-0.392 (0.183)
Dwelling (1 = Flat)	-0.029 (0.193)	0.105 (0.165)	-0.103 (0.180)
Ownership (1 = Rented)	-0.873*** (0.254)	-0.335* (0.185)	0.681*** (0.197)
Household Area	0.005*** (0.001)	-0.001** (0.001)	-0.005*** (0.001)
Travel Characteristics			
Distance to workplace	-0.005 (0.009)	-0.04*** (0.008)	0.073*** (0.008)
Travel frequency (1 = More than once)	-1.197*** (0.314)	-0.074 (0.196)	-1.913*** (0.531)
Built Environment Characteristics			
Population Density	0.000 (0.05E)	0.00 (0.00)	-0.001 (0.001)
HH distance to city centre (km)	0.014 (0.025)	0.040** (0.019)	0.081*** (0.021)
HH distance to district centre (km)	-0.134*** (0.048)	0.004 (0.038)	0.013 (0.042)
Workplace distance to city centre (km)	0.026** (0.011)	-0.007 (0.010)	-0.077*** (0.01)
HH distance to metro station (km)	0.044 (0.063)	-0.006 (0.042)	-0.16** (0.066)
HH Bus stops density	0.007 (0.007)	-0.009 (0.006)	0.006 (0.007)
Workplace dis. to metro station (1 = More than a km)	0.444*** (0.173)	0.214 (0.140)	-0.429*** (0.155)
Parking availability at workplace (1 = Yes)	0.615*** (0.183)	-0.20 (0.140)	0.366 (0.166)
Street Intersection Density (SID)	0.000 (0.000)	0.002 (0.000)	-0.001 (0.001)
Street width	0.064*** (0.014)	0.002 (0.012)	-0.042*** (0.015)
Footpath availability (1 = Yes)	0.770*** (0.178)	-0.397** (0.155)	-0.201 (0.177)
Traffic jam level in neighbourhood (1 = High)	0.085 (0.213)	-0.189 (0.168)	-0.164 (0.181)
Street cleanliness (1 = Unclean)	-0.150 (0.332)	-0.238 (0.223)	-0.130 (0.248)
Street safety (1 = Unsafe)	-0.602 (0.408)	-0.016 (0.248)	0.150 (0.267)
Travel Attitude			
Cost Minimising over Time Minimising	-0.746*** (0.223)	-0.0423** (0.165)	0.397** (0.024)
Comfort Maximising over Time Minimising	0.54*** (0.202)	-0.05 (0.172)	-0.589*** (0.214)
_cons	-6.406129 (0.891)	1.00*** (0.594)	-0.148 (0.685)
Number of obs	1,456	1,456	1,456
LR chi2(32)	616.9	284.36	307.89
Prob > chi2	0	0	0
Pseudo R2	0.386	0.198	0.242

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

We find that an increase in household distance to the city centre increases the likelihood of commuting by public transportation and two-wheelers but does not have any significant association with the usage of cars. This finding is in contrast to previous studies that show that commuters in the city's outer areas are more likely to rely on cars for commuting to the workplace (Ding & Cao, 2019). In Delhi, the neighbourhoods in the city's inner area are dominated by high-income households. On the other hand, households with low economic status are more dominant in the city's outer areas and periphery as compared to the city's inner areas. While the neighbourhoods in the city's outer areas are still developing their public transportation network, the high usage of public transportation in such areas by low-income households shows the importance of making the city's public transportation network more accessible in such areas to enhance the socio-economic development in the city.

Equally important is to reduce the overuse of cars in areas closer to the district centres. The model result shows that those who live near the district centre were found to rely more on cars for commuting to the workplace. As the district centres have high building density, the overuse of cars in such places may have negative implications on environmental sustainability and can lead to congested roads and discourage the use of non-motorised transportation (NMT). Planning interventions to reduce the use of cars and enhance the NMT such as pedestrian-friendly streets, separate bicycle lanes, and congestion pricing can be considered.

Considering the influence of street design, we find street width positively associates with the usage of cars and negatively associates with the usage of public transportation. This finding agrees with the general observation that neighbourhoods with wide streets do make it easier for commuters to use cars for commuting thereby discouraging the use of public transportation. We also find that footpath availability is positively correlated with car usage. While, previous studies have shown that the availability of footpaths enhances the walking culture and improves accessibility to transit stations, in Delhi we find contradictory results. A possible explanation for this is that in Delhi, neighbourhoods with footpath availability are more planned and thus, are dominated by high-income households. Thus, in such neighbourhoods, there is more likelihood of the usage of cars. Other street-related features, such as street traffic jams, street cleanliness, and street safety had no significant relationship with the mode choice.

We now examine the influence of two travel characteristics on commuting mode choice. First, is the commuting distance, increase in which decreases the likelihood of commuting by two-wheeler and increases the likelihood of commuting by public transport. This is an expected finding as reported in many previous studies that show the higher utility of commuting by public transport in long-distance commuting (Rasca & Saeed, 2022; Ko et al., 201). However, the model does not show any significant association between commuting distance with the likelihood of commuting by car. The second travel characteristic is travel frequency, which shows that commuters who visit their workplace more than once a day, have less likelihood to use both public transport and car. This hints that such commuters rely more on non-motorised form of transportation and works very near to their home.

With regard to the socio-economic and household characteristics, we find many of them have a strong and significant influence on commuters' choice of travel mode. With an increase in household income commuters are more likely to use cars for commuting to the workplace and less likely to rely on the use of two-wheelers and public transport. Many previous studies especially in the global south have found that income is a strong determinant of travel mode choice, which our study also confirms (Nakshi & Debnath, 2021; Nkeki & Asikhia, 2019). Gender has a strong influence in deciding the mode choice of commuters, where females are less likely to use cars and two-wheelers and more likely to rely on public transport for commuting to the workplace. Other characteristics such as age, education level, number of school-going children, household ownership, and household area also have significant influence over mode choice.

With regard to the travel attitude, we find respondents who have a higher preference for minimising commuting cost over minimising commuting time are more likely to use public transportation and less likely to use private transportation. On the other hand, those who prioritised maximising commuting comfort over minimising commuting time are more likely to use a car and less likely to use public transportation. One can easily relate to these findings as public transportation is often cheaper but less comfortable than private transportation which is costly but more comfortable.

The above analysis shows that commuting mode choice is influenced by different characteristics related to the socio-economic, built environment and travel attitude. While socio-economic characteristics such as income, education, and gender have a stronger influence on the mode choice, built environment characteristics, namely, household distance to city centre, district centre and metro station, along with workplace distance to the city centre and metro station also have a significant relationship with the commuting mode choice. Commuting distance to the workplace also influences the usage of public transport. The model results also show that commuters' travel attitude is a significant predictor of their mode choice, and the preference of cost, time and comfort are important considerations in choosing the travel mode.

The study findings on mode choice can be contextualised under the theory of planned behaviour (Ajzen, 1991). The theory suggests that attitude, social norms, and perceived behavioural control are important psychological factors that influence human behaviour. From our results, it is clear that the travel attitude of minimising cost and time is an important psychological factor that shapes mode choice behaviour. This can also be linked to the economic-rational choice framework as per which an individual acts to maximise his/her utility by minimising cost. The higher likelihood of using a car with an increase in income points to the existing social norm that links car usage with income. Less likelihood of females and aged group using a car also point to their inability to use a car due to a lack of driving skills and health and societal barriers, which is a perceived control behaviour. While the study does not consider a direct influence of the built environment on psychological factors, it should be acknowledged that the built environment can possibly influence the individual's attitude and perceived behavioural control which can affect their mode choice. In this manner, the theory of planned behaviour provides a conceptual framework to analyse how factors other than the built environment can play a determinant role in commuting mode choice behaviour.

A note on establishing causal inference.

Undertaking a regression analysis to study commuting mode choice with predictors including the built environment, travel attitude, and socio-economic factors presents a complex challenge in establishing causal inference. Regression analysis, in and of itself, does not establish causation. Instead, it identifies and quantifies associations or relationships between variables. While it can provide valuable insights into potential causal relationships, making causal claims based solely on regression results can be problematic. The regression analysis involves multiple independent variables, some of which may be subject to endogeneity. We use this section to briefly discuss the causal inference considerations, endogeneity, and methods to address it in this specific regression analysis.

Causal inference aims to determine whether changes in predictor variables (independent variables) cause changes in the outcome variable (commuting mode choice). Establishing causality is challenging because it requires addressing potential endogeneity and identifying exogenous variability. Endogeneity may arise if travel attitude is included as an independent variable. Travel attitude can be influenced by past mode choices, and at the same time, it can influence current mode choices. This creates a feedback loop that can lead to endogeneity. Socio-economic factors, such as income or employment status, can also be endogenous if they are influenced by past mode choices. For example, people who use a car may earn more because they have access to better job opportunities. To address endogeneity and enhance causal inference in regression analysis, the following methods can be utilised:

(i) Using Instrumental Variables (IV) technique to identify instruments that are correlated with the endogenous variables (e.g., travel attitude or socio-economic factors) but are not directly related to mode choice. For example, one could use distance to public transportation as an instrument for travel attitude, assuming that it affects travel attitude but is not directly related to mode choice.

(ii) Time-lagged Variables: Create lagged versions of endogenous variables (e.g., lagged travel attitude) to address potential reverse causality. Lagged variables can help capture the temporal order of causality.

(iii) Difference-in-Differences (DiD): If the data includes repeated observations for individuals over time, a DiD approach can be utilised. DiD compares changes in mode choice over time for individuals who experience changes in independent variables (e.g., changes in travel attitude) with those who do not.

(iv) Fixed Effects Models are also modelling technique models to control for unobserved heterogeneity that may be driving endogeneity. Fixed effects account for time-invariant characteristics of individuals that could be affecting both mode choice and the independent variables.

(v) Sensitivity analyses and robustness checks can also be used to assess the robustness of results to different model specifications and potential endogeneity.

Exogenous variability refers to variations in the independent variable(s) that are driven by factors external to the model, such as random shocks or policy changes. Exogenous variability is crucial for establishing causality because it helps isolate the causal effect of the independent variable from endogeneity.

To identify exogenous variability in the regression analysis, one may consider:

(i) Policy Changes to examine cases where external policy changes (e.g., introduction of public transportation improvements) introduce exogenous variability in the built environment or travel attitude.

(ii) Natural Experiments to look for natural scenarios, such as sudden disruptions or events that affect travel patterns and can be treated as exogenous shocks.

(iii) Randomized Experiments can also be conducted where specific interventions are randomly assigned to individuals or areas to generate exogenous variation.

The limitations of regression analysis in establishing causation have important implications for drawing main conclusions and making policy considerations. Given that regression analysis can only identify associations or correlations between variables, conclusions should be framed in terms of associations rather than causal relationships. It's important to acknowledge the limitations and uncertainties associated with causation when presenting results. Policymakers should be informed that while regression analysis can highlight relationships between variables, it does not provide direct evidence of causation. When formulating policies or making decisions based on regression results, policymakers should consider the possibility of confounding factors or reverse causation. Combining multiple methods and conducting thorough sensitivity analyses can help strengthen the causal inference in the analysis of commuting mode choice.

3.7 Conclusion

Studies examining the linkage between travel behaviour and built environment for residents in cities of the global south are limited in the literature primarily due to a lack of publicly available data on travel behaviour. Cities like Delhi, which is one of the most populous cities in the world have much to offer in improving our understanding of travel and built environment that has largely been from the cities of global north, especially Europe and North America. Looking at this caveat, the study aimed at understanding the commuting behaviour to the workplace of working residents in Delhi.

A household survey was conducted in Delhi where we interviewed 1679 working individuals, collecting information on their travel characteristics, built environment characteristics, socio-economic and household characteristics, travel attitude and residential location choices. Using regression modelling, we analysed the factors affecting commuting distance and mode choice. For commuting distance, the model results showed that those who commute longer distances to the workplace are more likely to live near the city centre and/or work at job locations lying away from the city centre. This signifies the possibility of low-job housing balance in the city's inner area. This is an important finding of this study that shows the outward mobility of commuters in Delhi in contrast to inward mobility which studies find in monocentric cities. To reduce commuting distance one policy recommendation can be to enhance the white-collar job opportunities near the city centre. However, as the city centre remains the city's prime location owing to its historic and cultural significance, enhancing the job opportunities may lead to traffic congestion and over-densification in the city's inner area and thus, any such move needs to be strategically planned and require further research.

The result also shows that those who live near metro stations commutes longer distance. This showcases the higher utility of using the metro for commuting longer distances. We also find commuters have a greater tendency to select the commuting distance based on their travel attitude of minimising commuting cost and maximising commuting comfort over minimising commuting time. While the study could not explore fully the causal mechanism behind travel attitude and commuting distance due to lack of data, the results do hint towards the role of residential self-selection as an underlying cause behind the association between commuting distance and the built environment.

Analysing the variation in travel mode choice, we found a higher likelihood of commuting by public transport among respondents who have workplaces near the city centre, while those who worked in the city's outer area are more likely to commute by car. Increasing the use of public transport for commuting to workplaces located in the city's outer areas will require improving the metro rail network connectivity at such locations. Improving accessibility to metro stations in city outer areas also becomes important as in the city such areas are mostly inhabited by low-income households. Thus, improving accessibility to metro stations can be a tool for socioeconomic development in the city. The study also notes some planning measures to discourage the overuse of cars in areas closer to the district centres.

Examining why the respondents prefer commuting by private mode over public mode, the majority of the survey respondents reported poor connectivity to transit stations as a primary cause that makes commuting by public mode less preferred over private mode. Some of the respondents who shifted to commuting by metro reported the operationalisation of a new metro station in their neighbourhood. Thus, findings from the model and descriptive data analysis highlight that by enhancing the accessibility to transit stations for commuters in Delhi the likelihood of commuting by metro can be increased. At the same time, we also find that for many respondents the use of public transport arises due to their preference to minimise travel cost over travel time, and due to their preference to live near transit stations. This suggests that there is an underlying causal mechanism linking travel attitude with residential location choice and commuting behaviour.

Based on the survey findings and result analysis, the study concludes that the built environment has a considerable influence on commuting behaviour. However, socioeconomic indicators, such as income and gender and importantly, as our study shows, travel attitudes have a stronger influence on commuting behaviour. This finding resonates with findings from other studies in the global south that seem to suggest that commuting behaviour in low-income countries is more influenced by the socio-economic characteristics of commuters and built environment has less degree of influence on commuting behaviour compared with cities in the global north (Nkeki & Asikhia, 2019). However, planning interventions related to the built environment as highlighted in this section can be thought of as policy measures to enhance access to the workplace and promote the use of public transport in Delhi.

While the study holds significance in terms of addressing the concerns of travel attitudes in examining commuting behaviour in cities of the global south, it has few limitations. First, the study does not address the influence of intra-household interaction on the travel behaviour of surveyed respondents. Second, the study does not model the non-linearity in explaining the causal mechanism between built environment and commuting behaviour which some studies in the recent past have shown. Third, the study although considers the residential location choice in the descriptive analysis, does not factor them into the statistical model due to the non-availability of data on households' preferences of residential location at the time of choosing their current residential location. Considering the complex nature of the relationship between the built environment and commuting behaviour, exploring the causality can be challenging. While there exist some casual relationships as shown in this study, exploring the one best causal structure will require daily travel activity data of households across different time intervals and the use of sophisticated modelling techniques to parametrise the commuting behaviour, which future studies should aim to pursue.

Chapter 4

Analysing Inequity in Accessibility to Services with Neighbourhood Location and Socio-Economic Characteristics in Delhi

Chapter Overview: This study addresses accessibility as a key driver of social equity and sustainability. While previous research has explored spatial variations in accessibility to services, limited spatial data at the neighbourhood level have hindered understanding of socio-economic inequities in accessibility in cities in the global south. To bridge this gap, we create a spatial database of 4,145 residential locations in Delhi, aggregating them into 1 km grid-shaped neighbourhoods. The neighbourhood's economic status is evaluated using a composite index of the built environment, land price, and household income collected through field surveys. Social characteristics are examined through the percentage of the scheduled caste (SC) population, considering their historical marginalization in Indian society. Using the E-2SFCA method, we calculate accessibility to four key services and employ the geographically weighted regression (GWR) model to explore inequities in accessibility based on neighbourhood location and socio-economic characteristics. Findings reveal inequity in accessibility to services at the neighbourhood level is primarily driven by spatial location rather than income or percentage of the SC population. Moreover, the influence of socio-economic characteristics on accessibility varies across locations. The study findings can help planners in Delhi in prioritising the distribution of critical services according to the neighbourhood's characteristics.

4.1 Introduction

Findings from previous studies show that cities across the world have spatial and social inequity in the distribution of services, whereby neighbourhoods characterised by low-income residents and socially marginalised communities have poor access to services (Zhao et al., 2020; Saroj et al., 2020). Much of our understanding of the relationship between accessibility to services and neighbourhood socio-economic characteristics has come from cities of the global north, fewer studies have explored the relationship between them for cities from the global south using the social equity perspective (Li et al., 2021). An important factor limiting such studies from the global south is the unavailability of spatial data related to the neighbourhood’s socio-economic and built-environment characteristics, and distribution of public services (Abascal et al., 2022). Moreover, the lack of neighbourhood maps poses a major hindrance to performing spatial analysis at an appropriate scale that can showcase the variation in accessibility to public services.

The study aims to understand the spatial inequity in accessibility and its relationship with neighbourhood characteristics in the city of Delhi, India. Delhi, like many other cities in the global south, lacks spatial data on the distribution of public services and neighbourhood socio-economic characteristics. To overcome the data constraints, we first create a spatial database of all the 4,145 residential locations in Delhi attributing every residential location with data on total population and scheduled caste population which serve as an indicator of social characteristics. Furthermore, we categorise the residential locations under different economic statuses using data on its built environment, land price and mean household income, where data on household income is collected through a field survey using a stratified sampling technique. Finally, the socio-economic characteristics of the residential locations are aggregated at the geographical scale of a 1 sq. km grid structure that serves as the neighbourhood map.

The necessity to make a grid-shaped neighbourhood as a spatial unit of analysis arises as Delhi lacks an administrative boundary map at the neighbourhoods or block level. While there exist ward-level maps in Delhi, wards as a spatial unit of analysis can be too large to produce accurate results, as we notice the population and services within the wards are not uniformly distributed. The use of larger administrative boundaries is a limitation that has been acknowledged in previous studies (Ashik et al., 2020). While studies indicate that data aggregation within grid cells yields superior outcomes compared to administrative units (Rothlisberger, 2017; Ahmed & Bramley, 2015), the selection of an appropriate spatial scale for analysis remains a contentious matter, contingent upon the study's objectives (Viegas et al., 2009). For this study, the utilization of a 1 sq. km grid cell is deemed optimal, as it effectively reveals spatial disparities in accessibility while accommodating multiple residential locations that contribute to heterogeneity in the neighbourhood's socio-economic characteristics.

Keeping with the study aim we examine the following questions – Does there exist spatial inequity in accessibility for different neighbourhoods in Delhi? If so, does the spatial inequity in accessibility exist on grounds of the neighbourhood's spatial location or socio-economic status? Using the E-2SFCA method we measure accessibility to 4 different services, namely schools, hospitals, entertainment facilities and jobs. Further, we build a geographically weighted regression model to examine the relationship between access to these four services with neighbourhood socio-economic characteristics using indicators of neighbourhood richness index and percentage of scheduled caste population. We also use neighbourhood population density, and its proximity to the city centre and district centre as a control variable.

The study becomes significant as it measures the accessibility for every residential location in Delhi, making it the most comprehensive study, best to our knowledge, on accessibility to services in a city in the global south. The choice of city Delhi is noteworthy, as Delhi has seen a massive increase in migrant population in the last few decades, making it one of the most populous cities across the world (UNDESA, 2018). With increasing urbanisation, the existing settlements have densified beyond their carrying capacity and new settlements in the city's peripheral areas have emerged (Naikoo et al., 2020). This has disturbed the demand-supply ratio for services and has adverse implications on individual well-being and the city's economic development and environmental sustainability. We believe findings from this study can throw light on the current distribution of public services in Delhi and can help city planners plan the distribution of services incorporating the ideals of equity in accessibility to have sustainable urban development.

The rest of the paper is organised as follows. Section 2 provides a brief review of the literature on inequity in accessibility to services with neighbourhood characteristics. In section 3, we provide details on the research statistical and spatial data. Section 4 provides the details of the accessibility measure and the spatial regression model. Section 5 provides study results and section 6 builds a discussion on study findings. Section 7 concludes the paper.

4.2 Literature Review

The importance of public services in enhancing the viability of urban life has been greatly emphasised in the literature (Allen & Farber,2020; Liang et al., 2020; Karji et al., 2019). Various studies have highlighted the positive impacts of access to different types of public services on various aspects of urban well-being. For example, access to green spaces has been shown to improve both physical and mental health (Wang et al., 2017), and access to jobs has been found to have a positive impact on neighbourhood median household income (Delmelle et al., 2021). Additionally, access to transit stations has been linked to improvements in environmental wellbeing (Basu & Ferreira, 2021). On the other hand, studies also show that inequitable distribution of resources can have adverse effects on individual and environmental wellbeing and may result in residential segregation in a city (Cortes, 2021; Galaskiewicz et al., 2021). As the spatial mismatch hypothesis (Kain, 1968) suggests, the inequitable distribution of public services results in a spatial mismatch where a group of the population enjoys better accessibility to opportunities than the other groups. Thus, to have equitable development, an understanding of resource allocation in a region becomes crucial.

Today, the understanding of how public services should be allocated has moved from the theory of even allocation to the theory of justice (Li et al., 2020). While having an equal spatial distribution of services in a city is ideal, the spatial distribution of resources should be such that it favours the most disadvantaged section of society (Rawls, 1999). In this context, equity is seen as a significant pillar of urban social sustainability. Social equity recognizes that certain disadvantaged groups may face greater obstacles to accessing resources and aims to eliminate these disparities by ensuring that resources are distributed fairly and in a way that benefits the most disadvantaged members of society (Meerow, 2019). As discussed in the review article by Dempsey et al. (2011), the concept of social equity at the neighbourhood level manifests itself in terms of fair access to resources by all. In a geographical sense, spatial inequity will exist if certain areas due to their spatial location are deprived of key resources.

To examine the existence of social and spatial inequity in access to resources studies have used accessibility as a standard measurement tool. However, they differ in terms of considered public services and the manner in which accessibility is measured and modelled with neighbourhood characteristics. For example, many studies have considered accessibility to parks/green spaces (Liu et al., 2021; Chang et al., 2019; Sharifi et al., 2021; Guo et al., 2019; Chen et al., 2020; Wu et al., 2020), while others have focussed on hospitals (Zhao et al., 2020; Jin et al., 2022; Mayaud et al., 2019), supermarkets (Li et al., 2019), public transport (Pereira et al., 2023), street infrastructure (Li et al., 2022).

With regard to the accessibility measures, studies in past have primarily relied on gravity-based methods (Chang et al., 2019; Sharifi et al., 2021) or the 2-step floating catchment area (2-SFCA) methods (Guo et al., 2019; Jin et al., 2022; Wu et al., 2020). Some studies have used distance/time-based measures (Chen et al., 2020; Cortes, 2021). While the gravity model or distance/time measures can provide a realistic measure of accessibility experienced by service users, they ignore the barrier in accessibility that may arise with an increase in the service demand potential. As the increase in urbanisation has resulted in a demand-supply mismatch it becomes crucial to incorporate the demand potential of a neighbourhood in measure of accessibility (Liang et al., 2023).

To incorporate the demand potential, studies have used the 2SFCA method, which assumes uniform accessibility for all points lying inside the catchment area. A more robust approach which we use in this study is the Enhanced 2-SFCA. The method developed by Luo & Qi (2009) overcomes the problems of uniform access within the catchment area by combining the demand potential of the 2SFCA method with the distance decay function of the gravity model. Many studies in the recent past have used this method to measure accessibility to services (Luo et al., 2020; Zhao et al., 2020)

Studies have modelled the variation in accessibility with neighbourhood characteristics using different modelling techniques which primarily include linear regression (Chang et al., 2019; Guo et al., 2019; Chen et al., 2020; Wu et al., 2020) and bivariate LISA analysis (Sharifi et al., 2021; Jin et al., 2022). Few studies have used correlation analysis (Li et al., 2019). We find the use of regression models like OLS in this context is insufficient due to the spatial correlation among the variables, which violates the homoscedasticity assumption of OLS. Furthermore, the OLS is a global model that provides non-spatial regression coefficients and, therefore, cannot demonstrate the heterogeneity in the relationship across space (Fotheringham et al., 2002).

To overcome this issue, this study uses one of the widely used spatial regression methods, geographically weighted regression (GWR) which shows the local variation in the relationship between accessibility and neighbourhood characteristics. Yang et al. (2022) find that while many previous studies have examined the inequity in accessibility to parks for different socio-economic groups, they have ignored the inequity in the relationship that may occur at a specific space in the study area. Neglecting the local variations can cancel out the opposite nature of correlations between the accessibility and socio-economic in specific parts of the study area and thus, can provide inaccurate results.

Analysing the variation in accessibility with neighbourhood socio-economic characteristics, numerous studies in the global North have examined the inequity in accessibility for social groups mainly based on race and ethnicity (Talen, 2022; Liu et al., 2021; Kwate et al., 2013). However, such studies from the global South are limited. The inclusion of socially marginalised communities becomes important to understand the extent of social inequity in access to services that go beyond economic inequality. Studies in the Indian context show that some of the caste and tribal groups, particularly scheduled caste (SC) and scheduled tribes (ST), in past have faced barriers in access to basic services, such as education and health, on account of historical discrimination (Raghavendra, 2020; Kale et al., 2022; Kusuma et al., 2018; Sundaram & Tendulkar, 2003). While these studies showcase the caste-based inequity, their findings tend to be at a more aggregated level, such for a city or village and do not factor in the spatial variations within a region. Except few, most of these studies relied on survey data that was not representative of the whole city or region and used non-spatial methods for data analysis, which makes the study results less accurate.

We also find studies report a lack of neighbourhood income data as a study limitation which they overcome using data on the neighbourhood’s land price or house rent (Zhao et al., 2020; Chen et al., 2020). Although this is an accepted approach in the literature, it can be made robust using observations from the field. In this study we combine the neighbourhood land price data with sampled data on household income collected from field surveys, to generate the neighbourhood richness index as a proxy of neighbourhood economic status.

Another important aspect related to neighbourhood characteristics is the spatial location. While relating neighbourhood socio-economic status with accessibility, we find only a few studies have factored the neighbourhood's locational characteristics, especially the proximity to the city centre as a control variable (Chang et al., 2019; Guo et al., 2019). Naess (2019) in their critique of the literature on land use and transportation, mentions that the location of a neighbourhood relative to the city centre or any other lower-order centre has an impact on the distribution of resources. As seen in many cities, there exists the urban–suburban contrast in the distribution of services, where neighbourhoods around the city centre are denser and have more demand potential as compared to neighbourhoods lying in the city's outer areas. This enhances the distribution of different services around the city's inner areas and increases accessibility for the residents living in such neighbourhoods. Thus, accounting for neighbourhood spatial location becomes important while analysing the impact of neighbourhood’s socio-economic characteristics on access to services.

4.3 Data

4.3.1 Spatial mapping of Residential Locations and Services

One significant difficulty that we encountered in the research was the unavailability of spatial data on the population of different localities in Delhi. While previous studies have spatially mapped the population at the ward level, we could not find any study that has performed a similar exercise for spatial units below the ward level in Delhi. Thus, to calculate the accessibility scores for all residential locations across Delhi, we first created a spatial database of residential locations with their population size. We accessed the population database of residential locations in Delhi, recently released by the Delhi State Election Commission (SEC, 2022). The databases listed 4145 residential localities under the 250 wards of the Delhi Municipal Corporation (DMC). The locations were geocoded using Google Maps and then spatially mapped in ArcGIS software. The residential locations have a population within a range of 100 to 73476, with a mean population of 4555.

To measure the population data at the grid neighbourhood level, we summed up the population of the residential locations lying inside every grid cell. The process was performed using the spatial join tool in ArcGIS software. Figure 4.1 shows the map of residential locations in Delhi aggregated under the grid neighbourhoods of an area of 1 sq. km.

In this study, we calculate the accessibility to 4 services: schools, hospitals, entertainment facilities and job clusters. High access to hospitals and schools has a significant impact on human development, as discussed in the United Nations - Sustainable Development Goal 3 which is to enhance good health and well-being, and Goal 4 which is to provide quality education. High accessibility to jobs positively impacts the income of the poor and reduces income inequality, a common finding in many previous studies. Literature also shows that high access to entertainment facilities enhances the quality of life of people (Li et al., 2021).

To spatially map the different services, we first obtained the location address of all the services from their respective government department that is responsible for managing those service, as mentioned in Table 4.1. The location address of these services is publicly available online on the department websites. The data was verified through the visits to the department authorities which assured the full data accuracy of the publicly available data, that was available at the time of our data collection in October 2022. The location address of the services was then geocoded using Google Maps to create the spatial database of these services in Delhi. The resulting geocoded addresses for each service were then mapped in ArcGIS using the point shapefile format.

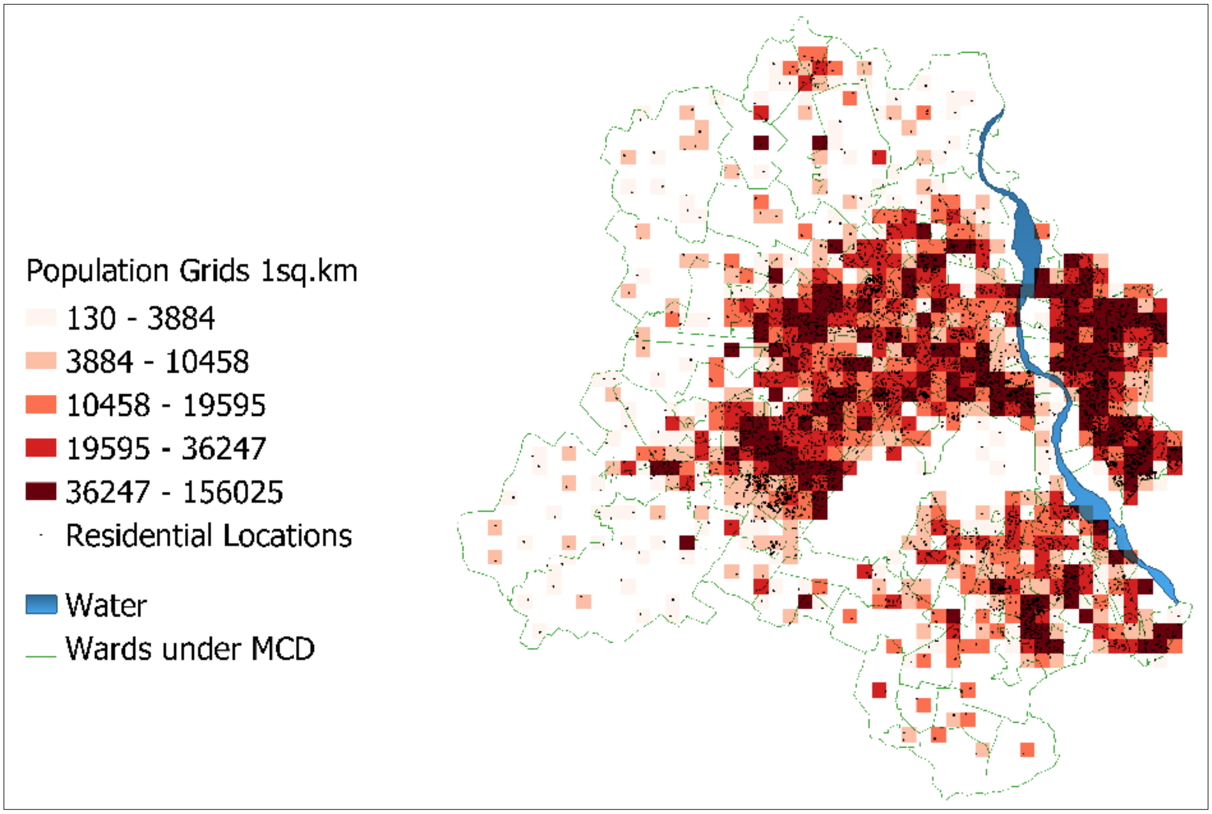

Figure 4.1: Neighbourhood locations in Delhi

Table 4.1: Type of services considered in the study.

Services	**Number of Observations**	**Variables**	**Source**
Schools	2403	Location of private and government schools of all levels.	Directorate of Education, Government of Delhi.
Hospitals	969	Location of public and private hospitals offering tertiary care	Department of Health and Family Welfares, Govt. of Delhi
Entertainment Facilities	827	Location of shopping malls, movie theatres, and registered restaurants	Department of Excise, Entertainment & Luxury Tax, Govt. of Delhi
Job Clusters	16	Location of main job clusters	Delhi Master Plan 2040

4.3.2 Neighbourhood Socio-economic Characteristics

The current research endeavours to evaluate the economic status of neighbourhoods by utilizing a composite index of the built environment, land price, and household income for each residential location. To begin with, we first characterise all the 4145 residential locations under three distinct elements of urban form, namely location compactness, building plot size, and street grid pattern. For the characterisation, we rely on the aerial imagery of these locations using the Google Earth software. This resulted in 8 categories of residential locations. The assigned attributes are based on the primary observation and should be considered only as a preliminary step to categorise the residential locations and to create a stratified sample of households.

From the 4145 residential locations, 200 locations were preliminarily chosen with equal representation across all 8 categories and distributed throughout the city. Within each selected location, 8-10 households were randomly chosen for the survey, with a minimum distance of 100 meters between them. In this manner, we collected household income data from a sample of 1700 households lying across the 8 categories of residential locations. The survey was completed with the help of 15 field surveyors from October to December 2021. During the survey, ethical guidelines and data collection norms were followed. Respondents were informed about the survey's purpose and only those who voluntarily agreed to participate provided written consent.

Our survey data showed that the households in residential locations categorised as A and B had a household income of more than $1333 per month. Households in residential locations C, D and E had a household income in the range of $1333 to $800 per month. Households in residential location F had a household income in the range of $800 to $400 per month. Finally, households in locations G and H had a household income of less than $400 per month. In this manner, we could classify the residential locations under 4 categories of economic status, namely high income (HI), upper middle income (UMI), lower middle income (LMI) and low income (LI). As the residential land price of a location correlates with its mean household income, we verified the assigned economic status of the residential locations using the residential land price data in those locations. We found that the economic status of residential locations closely correlated with their residential land price, showing the high accuracy of surveyed data. The residential land price database of different residential locations in Delhi was freely accessed online from different real estate property developers' websites. The summary of residential locations' built-up form, economic status and land price is presented in Table 4.2.

Table 4.2: Residential locations' built-up form and the assigned economic status.

Residential Locations Categories	Housing Compactness	Plot Area / Building type	Street Grid Pattern	Land Price (per sq. km)	Household Income (per month)	Economic Status
A	Low	Big plots	Regular	$10,320- $3280	More than $1333	High Income (HI)
B	Medium	Big plots	Regular			
C	High	Big plots	Regular	$2000-$1333	$1333-$800	Upper Middle Income (UMI)
D	Medium	Small plots	Regular			
E	Medium	High rise flats	Regular			
F	High	Small plots	Regular	$755	$800-$400	Lower Middle Income (LMI)
G	High	Small plots	Irregular	$533-$266	$400-$200	Low Income (LI)
H	Low	Small plots	Regular/Irregular			

After characterising the residential locations under the four economic statuses, we then look for residential locations lying under every grid cell. The study defines neighbourhoods as a grid-shaped area of 1 sq. km that may cover more than one residential location characterised by its total population and economic status. For a neighbourhood, the percentage of the population under an economic status i, denoted as $Population_i$, is calculated as the percentage of the total population under the economic status i to the total population of all residential locations lying in that grid cell. Thereafter, the neighbourhood richness index (NRI) is calculated as a weighted sum of the $Population_i$, as shown in Equation 4.1.

$$NRI = \sum (W_i * Population_i) \tag{4.1}$$

Where, W_i is the weight assigned to the economic statuses with the chosen weights 1, 0.66, 0.33, and 0 for economic status HI, UMI, LMI and LI, respectively. The selection of weights for the four economic status categories was based on a deliberate consideration of the degree of resemblance that each economic status bears to neighbourhood richness, while simultaneously maintaining equal intervals between the categories. The weights are relative scores, and any other weighing criteria will not change the nature of the results. NRI varies from 0 to 100, where a score of 0 represents neighbourhoods with residential locations of all low-income categories, and 100 represents neighbourhoods with residential locations of all high-income categories.

To determine neighbourhood social characteristics, the study uses the data on the percentage of the scheduled caste (SC) population for each residential location and consolidates it at the neighbourhood level. The data on the scheduled caste population is obtained from the population database of residential locations (SEC, 2022). Figure 4.2 shows the flowchart of the data preparation process.

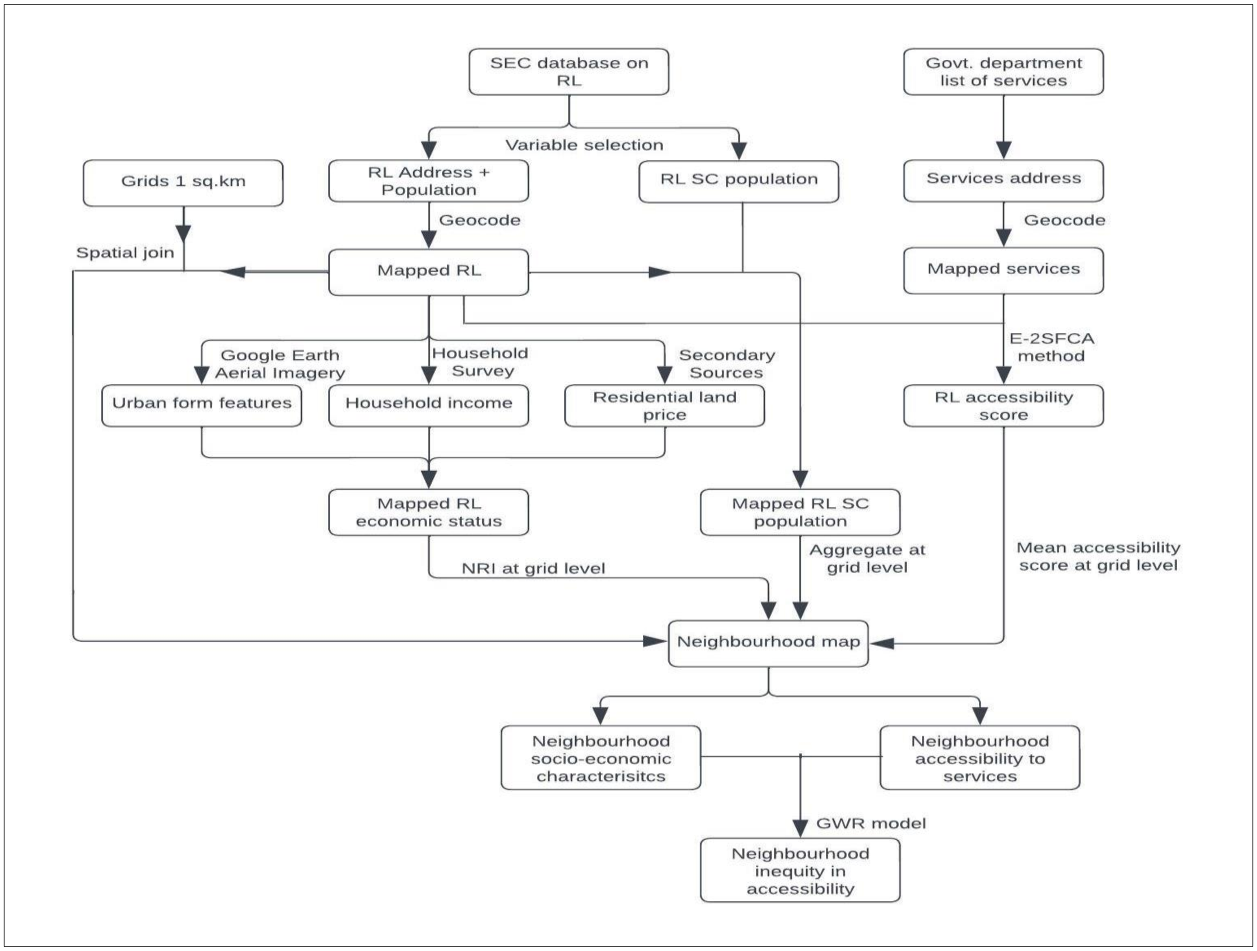

Figure 4.2: Flowchart of data preparation process

4.4 Methods

4.4.1 Accessibility calculation using the E-2SFCA method

In the literature, we find studies have used various measures of accessibility which can be categorised under location-based, infrastructure-based and person-based measures (Marwal & Silva, 2022). This study uses the Enhanced-2 Step Floating Catchment Area (E-2SFCA) method that comes under the Floating Catchment Area (FCA) modelling framework. The FCA method measures accessibility for a location in terms of the services available per unit population in a buffer area of that location. However, as the consumption of services can be from outside the buffer region, the demand for services in the FCA method is underestimated. To overcome this limitation, Radke and Mu (2000) proposed a spatial decomposition method called 2-step FCA (2SFCA). The method in the first step calculates the supply-to-demand ratio in a buffer region for every service. In the second step, accessibility for a location is calculated as the sum of supply to demand of every service that falls within the buffer region of the location. One major limitation of the method is that it is a binary construct, whereby it assumes equal accessibility to all the services that fall within the buffer region of the location and zero accessibility to those that fall outside the buffer region.

To overcome the limitations of the 2SFCA method, studies have used different impedance functions such as kernel density (Guagliardo, 2004) and Gaussian (Alford et al., 2008). In this paper, we use the E2SFCA method which has a step-up distance decay impedance function that provides different weights to different travel time zones within a buffer region of a location. The services and residential locations are referred to as supply locations j and the demand locations i, respectively. The method is executed as follows.

In the first step, we look for all the demand locations i that fall within the threshold travel time zone t from the supply location j. To account for the distance decay between supply and demand locations, we use the impedance function W_t which provides weights as per the travel time zone as shown in Table 4.3.

Table 4.3: Buffer thresholds and weights.

Buffer (t_r)	**Time zone**	**Distance zone**	**Weight (W_t)**
t_1	0-10 minutes	0-1 km	1
t_2	10-20 minutes	1-2 km	0.75
t_3	20-30 minutes	2-5 km	0.5

We chose the distance threshold based on the desired commuting distance range for different travel modes, commonly used in the literature which is 0-1 km for walking, 1-2 km for biking and 2-5 km for driving (Shen et al., 2017; Wu et al., 2020). Weights are assigned keeping the highest priority to walking mode and the lowest to driving mode.

The service-to-population ratio R_j is then computed for every service as shown in Equation (4.2).

$$R_j = \frac{S_j}{\sum_{i\in\{d_{ij\in t_r}\}} P_i W_t} \tag{4.2}$$

$$= \frac{S_j}{\sum_{i\in\{d_{ij}\in t_1\}} P_i W_1 + \sum_{i\in\{d_{ij}\in t_2\}} P_i W_2 + \sum_{i\in\{d_{ij}\in t_3\}} P_i W_3}$$

S_j denotes the capacity or supply potential of service j, P_i represents the population in neighbourhood i falling within the buffer region of service j such that ($d_{ij} \in t_r$), d_{ij} is the travel time between demand location i and supply location j, and t_r is the r^{th} travel-time zone (r ∈|1,2, 3|) within the buffer region.

In the second step, accessibility for every demand location i is calculated by summing up the supply-to-demand ratio R_j of all the services falling in the buffer time zone t_r from the demand location i and weighted by the distance decay function (W_t), as shown in Equation (4.3).

$$A_i = \sum_{j\in\{d_{ij}\in t_r\}} R_j W_t \tag{4.3}$$

$$= \sum_{j\in\{d_{ij}\in t_1\}} R_j W_1 + \sum_{j\in\{d_{ij}\in t_2\}} R_j W_2 + \sum_{j\in\{d_{ij}\in t_3\}} R_j W_3$$

A_i represents the accessibility for the population in neighbourhood i, R_j is the supply-to-demand ratio at supply location j that falls within the buffer region of neighbourhood i such that ($d_{ij} \in t_r$). Due to the unavailability of data regarding the supply potential of different service locations, we assume every service location has equal supply capacity.

4.4.2 Modelling variation in accessibility

To analyse the variation in accessibility, we first use the OLS method which also serves as our base model. The OLS is a global model that calculates the regression coefficients for the entire study area. This model can be useful when the relationship between dependent and independent variables is non-spatial in nature, i.e., not determined by the location. Since we are dealing with neighbourhood data, there likely exists spatial autocorrelation in the data. We check the spatial autocorrelation using Moran's I indicator, which shows how effectively the data points are clustered. To account for the spatial autocorrelation and the non-stationarity in the relationship between the dependent and independent variables we use the GWR model, as described below.

GWR was first introduced to the geography literature by Brudson et al. (1996) to capture the parametric non-stationarity in the regression models. In the OLS model, the regression coefficients are considered global or fixed while in the GWR model, they are estimated at each data location. While traditional regression emphasized on curve fitting or estimating the dependent variable, GWR is more about conducting inference on a spatially varying relationship (Páez and Wheeler 2009).

The basic form of the GW regression model is:

$$y_i = \beta_{i0} + \sum_{k=1}^{m} \beta_{ik} x_{ik} + \varepsilon_i \tag{4.4}$$

y_i is the dependent variable at location i; x_{ik} is the value of the k^{th} independent variable at location i; m is the number of independent variables; β_{i0} is the intercept parameter at location i; β_{ik} is the local regression coefficient for the k^{th} independent variable at location i; and ε_i is the random error at location i. The model measures the inherent relationships around each regression point i, where each set of regression coefficients is estimated by a weighted least squares approach.

We run the OLS and GWR models to analyse variation in accessibility for all four services, considered in the study. To run the model we use the 'GWmodel' package in R software (Gollini et al., 2015).

Model calibration

The GWR model requires specifying the spatial weighting matrix that sets the spatial relationship between the variables (Fotheringham et al. 2002). To build the weighting matrix three key elements are required: (i) the type of distance, (ii) the kernel function and (iii) its bandwidth. The GWR model allows for the use of different distance measurement techniques and kernel functions such as Gaussian, Exponential, Box-car, Bi-square and Tri-cube to define the weight matrix (Yu & Peng, 2019). In this paper, we selected the Gaussian function to determine the spatial weight. It is a continuous monotone decreasing function of the distance between the observation and calibration point. The weights are highest for observation at the model calibration point and decrease according to the Gaussian curve as the distance between the observation and calibration points increases. The Gaussian function is expressed as follows:

$$W_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right] \tag{4.5}$$

where d_{ij} is the distance between location i and j, b is the bandwidth. The key parameter in the kernel function is the bandwidth. Bandwidth can be specified either as the fixed distance or as a fixed number of local data (i.e., adaptive distance) (Gollini et al., 2015). Adaptive bandwidth is useful for data points that do not represent a regular configuration. It corresponds to the number of nearest neighbours and accounts for the local variations for each local calibration in the GW model. While the bandwidth can be user specified, we use the optimal kernel bandwidth function which finds the optimal bandwidth by minimising the model AICc value. AICc is the model fit diagnostic that takes into account the model prediction accuracy and complexity as a function of sample size (Akaike 1973; Hurvich, Simonoff & Tsai, 1998). Using the 'GWmodel' package in R, the optimal adaptive bandwidth for different accessibility models is calculated.

Further, to identify the best independent variables subset for the GW regression the study uses a model selection function using the 'GWmodel' package in R. The function first regresses all possible bivariate GW regressions using the independent variables and the dependent variable. The model which gives minimum AICc is retained. Table 4.4 shows the summary of variables used in the study.

Table 4.4: Variables summary

Variable Type	**Variable Name**	**Method**	**Min**	**Max**	**SD**
Dependent	Accessibility to Hospitals	E-2SFCA	0	11.88	5.54
	Accessibility to Schools		6.53	48.30	15.49
	Accessibility to Ent. Facilities		0	32.00	5.43
	Accessibility to Job Clusters		0	0.72	0.13
Independent	Distance from the city centre (km) (DCC)	Proximity Analysis in QGIS	0.68	35.76	14.28
	Distance from district centre (km) (DDC)		0.32	22.65	6.16
	Population Density	Total Neighbourhood Population / Neighbourhood Area	130	156025	21685.21
	Neighbourhood Richness Index (NRI)	Equation (4.1)	0	100	44.84
	Schedule Caste Population (%) (SC)	Total Neighbourhood SC Population / Total Neighbourhood Population	0	86.76	15.36

4.5 Results

4.5.1 Spatial variation in accessibility

Figure 4.3 displays the spatial variation in accessibility to different services. The analysis indicates that residential locations with the highest accessibility scores to hospitals are concentrated around the central west and southern parts of Delhi. Conversely, those residing in the periphery of Delhi exhibit the lowest accessibility to hospitals. Concerning schools, high accessibility scores are observed in residential locations in west and east Delhi, as well as in isolated residential pockets on the west and northwest boundaries of Delhi. This finding is somewhat surprising given that these areas are still in the development phase and possess semi-urban characteristics. Equally unexpected is the low accessibility to schools in the central part of Delhi.

Regarding entertainment facilities, the analysis identifies a substantial cluster of residential locations in south Delhi and a smaller cluster in central Delhi with high accessibility scores. In contrast, the outer part of Delhi and the peripheral region exhibit poor accessibility scores to entertainment facilities. Lastly, the analysis indicates that residential locations in south Delhi enjoy high accessibility to job clusters. Additionally, residential locations in southeast and central Delhi also exhibit high accessibility to job clusters, while those in the northeast and west Delhi display poor accessibility.

The results presented above demonstrate that there exists significant spatial variation in accessibility in the case of all the considered services, which can be defined as spatial inequity in accessibility. Notably, except for schools, residential locations situated in south and central Delhi exhibit the highest accessibility to services. Apart from high accessibility to services, we also find that these places provide better quality of services especially related to workplace and entertainment facilities. One can find some of the best shopping brands stores, food outlets, super speciality hospitals and offices of some of the top multinational companies in India in south Delhi.

In the last twenty years, Gurugram a city adjoining Delhi, has emerged as one of the best job locations in India. The proximity of Gurugram to south Delhi has made the south Delhi region a preferred residential choice for many high-income households which has led to the rise of residential flats and gated communities on the periphery of south Delhi (Bernroider, 2015). South and southwest Delhi has seen the highest rate of urban sprawl in Delhi during the period 2010-2020 (Sharma & Abhay, 2022). In contrast to south Delhi, the north, and northeast Delhi exhibit low accessibility to services. As one of the implications of low accessibility, these regions have lower land prices which has increased the population density of low-income households in the region. Also, people from north-east and south-east Delhi tend to commute more to the nearby town of Noida to access good quality services. An increase in commuting trips is one of the important factors behind the increase in vehicular emissions increasing air pollution (Sindhwani et al., 2015). Thus, we find that inequity in accessibility to services has an ever-lasting impact on household's residential location choice and commuting pattern in Delhi and its adjoining areas.

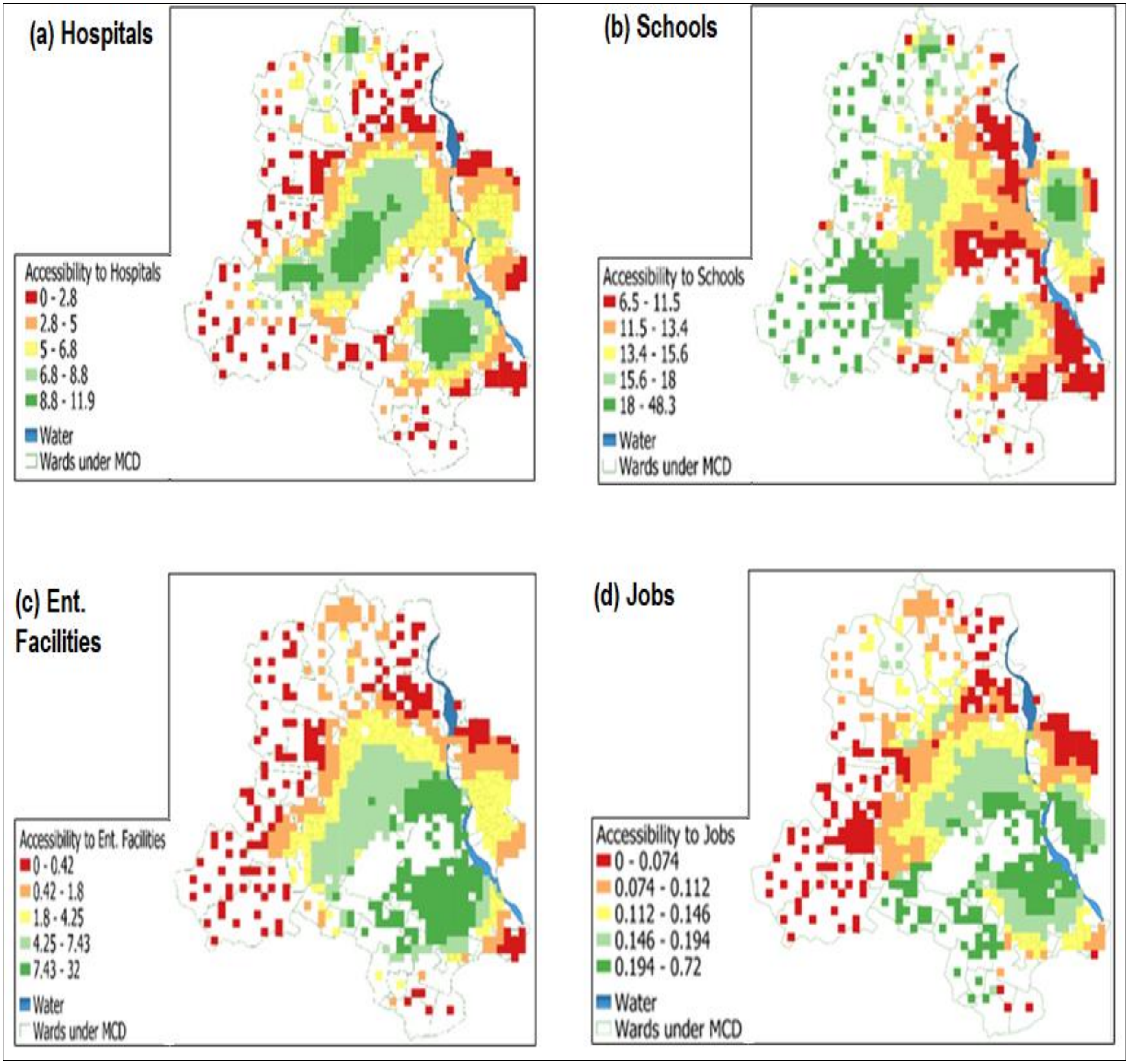

Figure 4.3: Spatial variation in accessibility to different services

4.5.2 Accessibility variation with neighbourhood characteristics

Initially, we conducted a pairwise collinearity test among the independent variables using the Pearson correlation coefficient test. We utilize a threshold value of 0.7 to assess the correlation between the variables (Wheeler & Tiefelsdorf, 2005). The results indicate a high correlation (>0.7) between the distance to the city centre (DCC) and the distance to the district centre (DDC) (as shown in Table 4.5). Given the local significance of distance to the district centre in urban planning, we exclude distance to the city centre from the model.

Table 4.5: Result of Pearson's correlation test

	DCC	**DDC**	**SC**	**NRI**	**PD**
DCC	1				
DDC	**0.825**	1			
SC	0.089	0.198	1		
NRI	-0.519	-0.489	-0.246	1	
PD	-0.299	-0.310	0.031	0.094	1

We proceed to conduct OLS regression to evaluate the elasticities at a global level. To ensure homogeneity in the measuring scale, we normalize the variables using the min-max normalisation technique which is a common normalisation technique that normalises features to a common range (0,1) while preserving the original data interpretability. The results of the OLS regression as shown in Table 4.6, demonstrate that the relationship between accessibility and all the considered neighbourhood characteristics is statistically significant. However, the magnitude and nature of impact vary in different accessibility models. For instance, distance to the district centre exhibits a strong correlation with accessibility to hospitals and schools, but the nature of impact differs.

Moreover, we examine the multicollinearity assumption using variance inflation (VIF), which indicates a value of less than 10 for every independent variable. This implies that there is no multicollinearity among the variables, and the variable selection is appropriate. The findings from the OLS model mostly conform to the findings from previous studies.

To capture the spatial heterogeneity in the relationship between the dependent and independent variables, we employ the GWR model, which builds on the OLS model. We aim to account for the spatial clustering of the accessibility variables, which is apparent from Moran's I value for accessibility to services, as shown in Table 4.7. A value of $I > 0$ denotes a positive correlation, i.e., similar values exist together. A value of $I < 0$ denotes a negative correlation, i.e., dissimilar values exist together. The value of $I = 0$

denotes no clustering and data is randomly distributed. The high degree of spatial clustering highlights the need to use the GWR model.

Table 4.6: Result of OLS regression

Variables	Hospitals	Schools	Ent. Facilities	Job Clusters	VIF
	b/se	b/se	b/se	b/se	
Distance to the district centre	-0.377***	0.418***	-0.074*	-0.162***	2.15
	(0.05)	(0.03)	(0.05)	(0.03)	
SC Population (%)	-0.139**	-0.054*	0.153**	0.020*	1.07
	(0.05)	(0.03)	(0.05)	(0.04)	
Population Density	0.020*	0.032*	-0.209***	-0.079*	1.14
	(0.05)	(0.03)	(0.05)	(0.04)	
NRI	0.182***	0.051**	0.346***	0.054**	1.99
	(0.03)	(0.02)	(0.02)	(0.02)	
Constant	0.505***	0.088***	0.036	0.219***	
	(0.03)	(0.02)	(0.03)	(0.02)	
R-sqr	0.348	0.274	0.375	0.118	
dfres	751	751	751	751	
AIC	-364.88	-1125.98	-491.41	-955.28	

* p<0.05, **p<0.01, ***p<0.001

Table 4.7: Moran's I value for accessibility to different services.

Accessibility	Moran's I
Hospitals	0.82***
Schools	0.88***
Ent. Facilities	0.92***
Job Clusters	0.79***

***p Value < 0.001

The weight matrix is formulated using the Gaussian kernel function for each of the four models with an optimal bandwidth score of 20. The GWR model output is presented in Table 4.8. Model fitness is evaluated using the adjusted R-square and AICc values, indicating that all the GWR models exhibit high accuracy and surpass the OLS models. Moreover, the F1 statistics presented in Table 4.9 suggest that the null hypothesis is rejected, demonstrating the existence of a significant difference between OLS and GWR models (Hu et al., 2018). To assess non-stationarity in the variables, F3 statistics are computed in the GWR models, as exhibited in Table 4.9. A statistically significant F3 value indicates that the relationship between the dependent and independent variables differs across space (Leung et al., 2000).

Table 4.8: Result of GWR models

Variable	Model 1: Accessibility to Hospitals			Model 2: Accessibility to Schools		
	Min	**Median**	**Max**	**Min**	**Median**	**Max**
Dist. To District Center	-2.56	-1.19	1.561	-0.608	0.024	0.739
SC Population (%)	-0.435	-0.087	0.547	-0.223	-0.007	0.136
Population Density	-0.258	0.034	1.193	-0.268	0.034	0.203
NRI	-0.208	0.073	0.705	-0.132	0.025	0.201
Intercept	-0.456	0.699	1.112	0.003	0.153	0.429
AICc	-1316.107			-1948.303		
R-square	0.852			0.805		
Adj. R-square	0.826			0.771		
Adaptive bandwidth	20			20		
Variable	**Model 3: Accessibility to Ent. Facilities**			**Model 4: Accessibility to Job Clusters**		
	Min	**Median**	**Max**	**Min**	**Median**	**Max**
Dist. To District Center (DDC)	-1.768	-0.181	1.806	-0.873	-0.159	1.604
SC Population (%) (SC)	-0.378	0.014	0.502	-0.529	0.001	0.274
Population Density (PD)	-0.565	-0.02	0.362	-0.147	-0.015	0.57
NRI	-0.077	0.097	0.813	-0.326	0.035	0.132
Intercept	-0.206	0.113	0.611	-0.160	0.194	0.453
AICc	-1718.403			-1800.49		
R-square	0.901			0.77		
Adj. R-square	0.884			0.73		
Adaptive bandwidth	20			20		

Table 4.9: Computed F statistics in the GWR models

F-Statistics	Variable	Model 1	Model 2	Model 3	Model 4
F3	Intercept	24.57***	6.33***	8.58***	3.99***
	Dist. To District Center	22.47***	8.56***	20.80***	10.42***
	SC Population (%)	1.38***	1.06	3.10***	1.02
	Population Density	4.58***	1.08	2.16***	1.32**
	NRI	11.90***	4.95***	20.83***	2.72***
F1	/	0.26***	0.31***	0.18***	0.30***
P-Value < ***0.001, **<0.01, *<0.05					

Figure 4.4(a) to (d) illustrates the spatial variation in the variables' coefficients as determined by the GWR models. Consistent with the findings of F statistics, Figure 4.4 reveals that most of the variables have a non-stationary effect on accessibility.

We now discuss the impact of neighbourhood characteristics on accessibility to all four services. Starting with the variable, distance from the district centre (DDC), as demonstrated in Figure 4.4(a) its impact on accessibility to different services, except for schools, is largely negative across the studied area. This suggests that accessibility decreases with increasing distance to the district centre in most of the examined regions. Previous research has also found a comparable correlation between accessibility and proximity to city or district centres. However, concerning accessibility to schools, the effect varies from negative in the east, southeast, and central north to positive in the rest of the area.

The variable SC in our study denotes the percentage of people belonging to the Scheduled Caste group. This particular segment of society has historically faced marginalization and challenges in accessing basic services. Our findings reveal that SC (as shown in Figure 4.4(b)) has a predominantly negative impact on accessibility to hospitals and schools across most parts of the region. This implies that areas with a high proportion of Scheduled Caste population tend to have low accessibility to hospitals and schools. However, in the case of accessibility to entertainment facilities and jobs, the influence varies across the region. Specifically, areas with a high percentage of Scheduled Caste population exhibit high accessibility to entertainment facilities and jobs in some parts of the region, while low accessibility in other parts.

Regarding the variable population density, as depicted in Figure 4.4(c), its influence on accessibility to various services is highly variable across the studied region. It has been observed that in some areas, an increase in population density leads to an increase in accessibility to services, whereas in other parts, accessibility declines with an increase in population density. Although previous research has indicated that dense areas generally have better access to services due to their high demand potential (Shi et al., 2020; Mayaud et al., 2019), our findings highlight that this may not always hold when considering local variations. One possible reason for this can be, in Delhi, we find the high-density residential neighbourhoods are too dense to have any vacant land space for a commercial service to locate which makes the service providers locate away from high-density neighbourhoods and thus, accessibility declines in such high-density areas. Also, as we found in our preliminary analysis, high-income households in Delhi are less likely to be found in high-density neighbourhoods this may incentivise the service providers to locate in low to moderate-density areas to gain from the high purchasing power of high-income households in such areas.

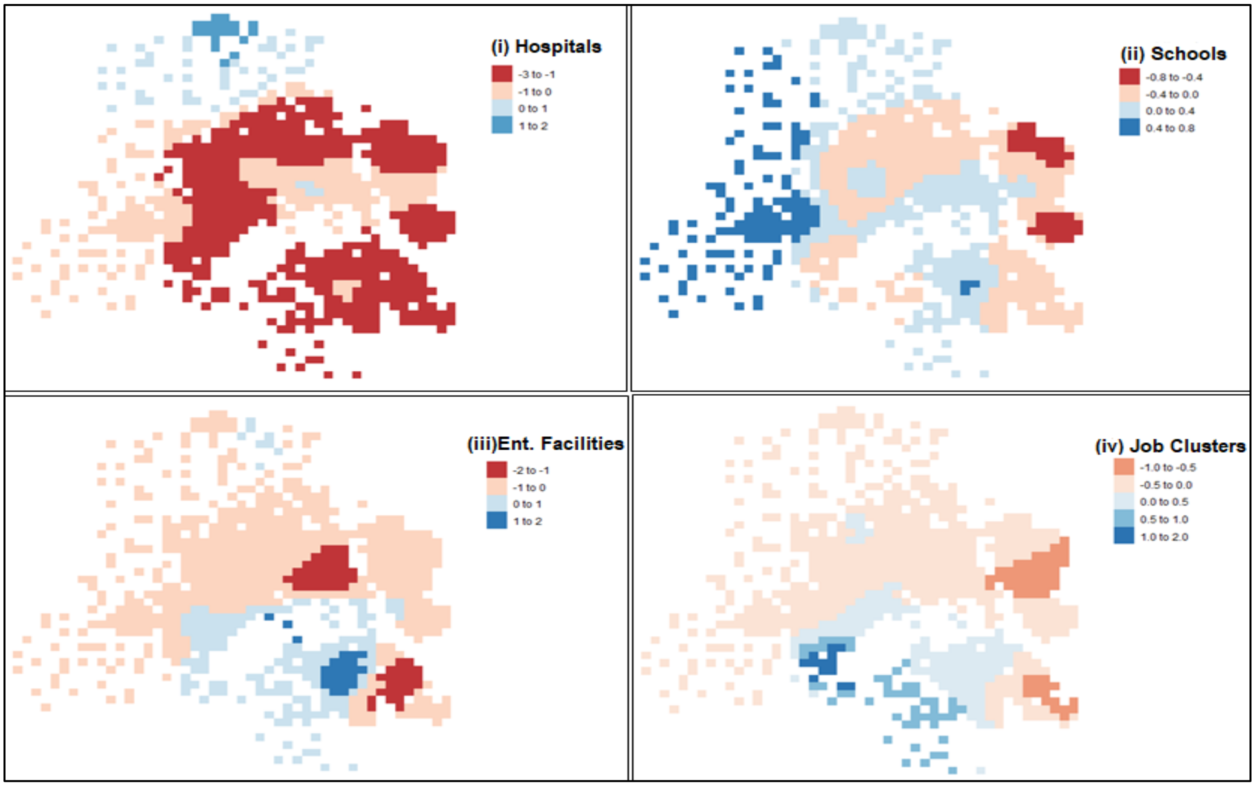

Figure 4.4(a): Relationship between neighbourhood distance from district centre and accessibility to different services

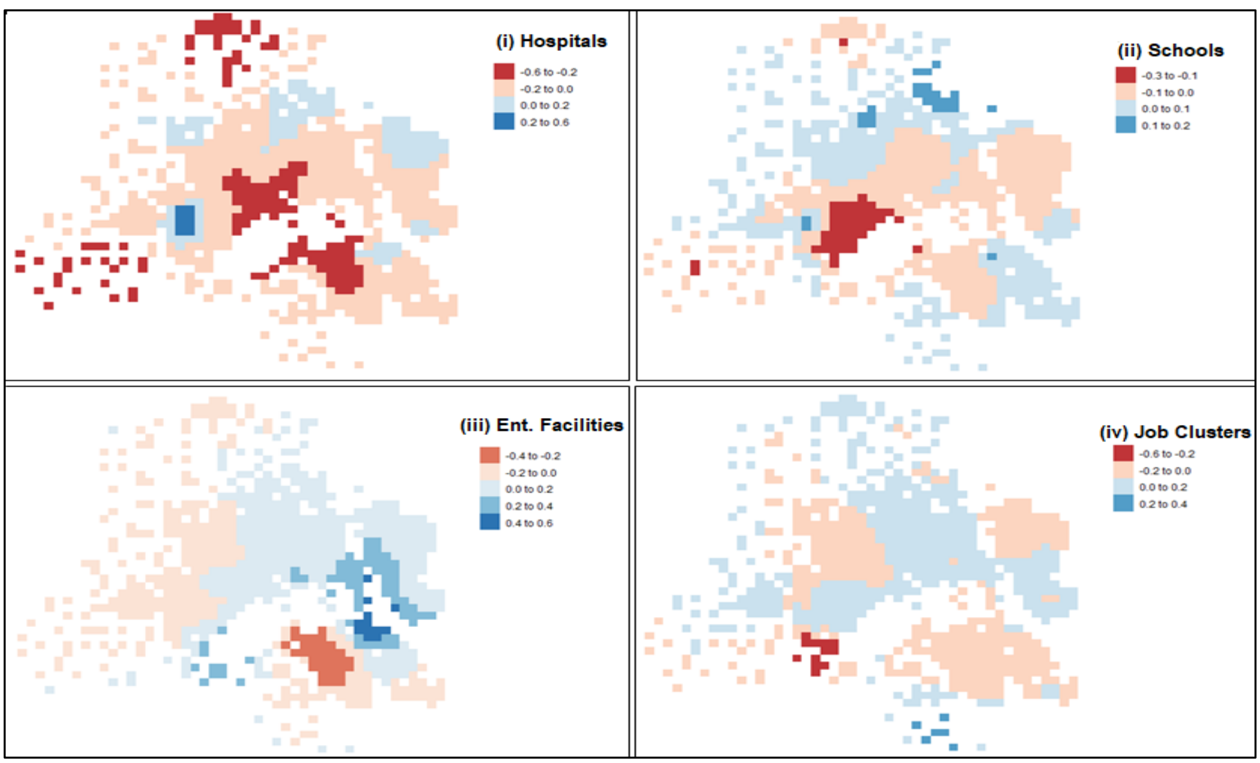

Figure 4.4(b): Relationship between neighbourhood SC population and accessibility to different services

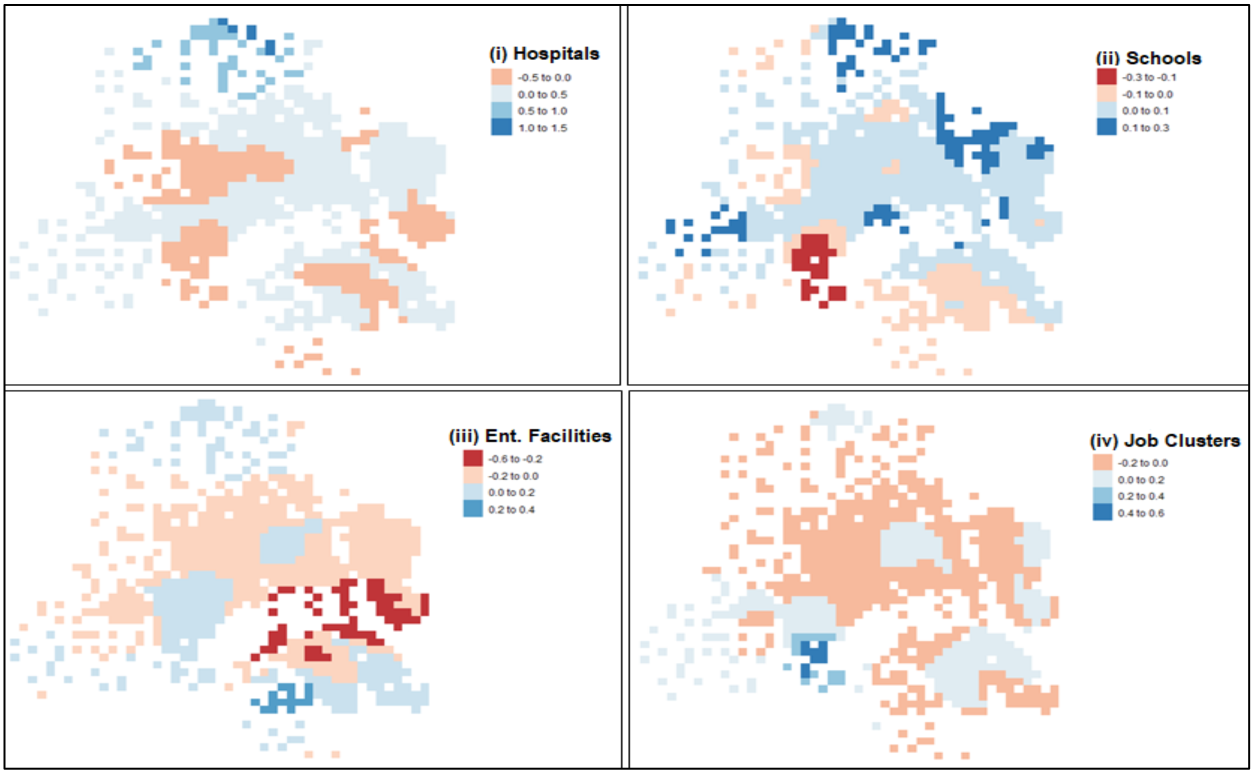

Figure 4.4(c): Relationship between neighbourhood population density and accessibility to different services

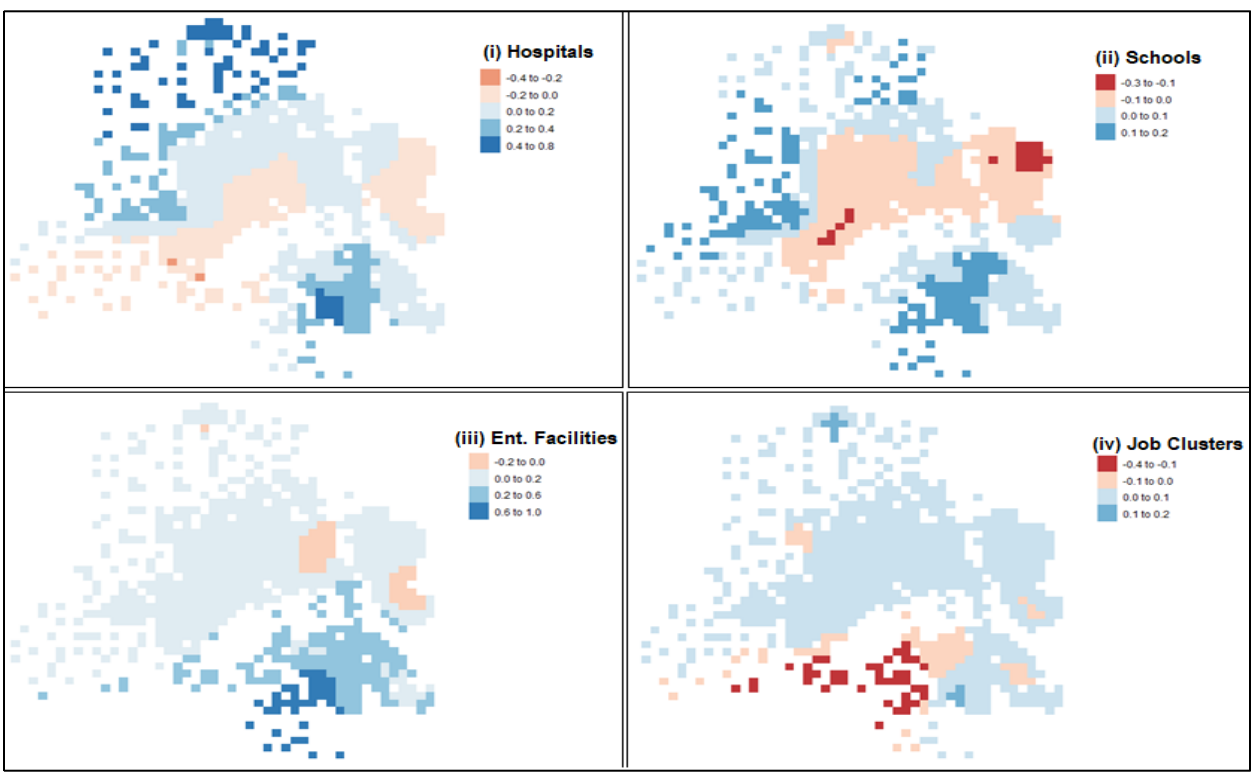

Figure 4.4(d): Relationship between variable NRI and accessibility to different services

In the final analysis of our study, we investigate the impact of the variable NRI on accessibility to different services, as depicted in Figure 4.4(d). Our findings reveal that NRI has a positive influence on accessibility to entertainment facilities and job clusters, indicating that richer areas are better served in terms of these amenities. This result is consistent with prior research in the field, and highlights the issue of vertical inequity in accessibility, as discussed in the literature (Litman, 2002). However, we also observe a mixed influence of NRI on accessibility to hospitals and schools, suggesting that accessibility to these services may decline with an increase in neighbourhood richness in some parts of the region. Thus, our study provides evidence that not all low-income areas necessarily experience poor accessibility to schools and hospitals.

4.6 Discussion

The limited availability of spatial data in cities of the global South, while presenting a significant challenge, also provides an opportunity to pioneer novel techniques in spatial data mapping and analysis. A predominant approach in previous studies within the global South has involved the utilization of administrative boundary units, such as wards or census tracts, in conjunction with secondary data sources to scrutinize disparities in accessibility. In contrast, our study endeavours to demonstrate the application of software tools like Google Earth and Geographic Information Systems (GIS) for the spatial mapping of services, the categorization of residential locations, and the intricate analysis of spatial data at the neighbourhood level. Furthermore, our approach integrates this spatial mapping with households' socio-economic data acquired through fieldwork, thereby augmenting the comprehensiveness of our analysis.

As cities in the global South continue to undergo rapid urbanization, attracting migrants from diverse socio-economic backgrounds, they inevitably witness the emergence of a multiplicity of settlement types. For instance, in the context of Delhi, we observe the coexistence of residential settlements such as slums and small-scale informal settlements alongside planned developments, each exhibiting distinct socio-economic characteristics when compared to their adjacent counterparts. In such instances, urban planners are compelled to employ advanced spatial data tools and techniques that can effectively delineate and differentiate the characteristics of these heterogeneous neighbourhoods. This becomes especially pertinent given that the nuances of diverse neighbourhoods tend to be overlooked when adopting a broader geographical scale of analysis, such as wards.

The advent of advanced Geographic Information Systems (GIS) tools and satellite data has rendered the high-resolution mapping of the physical attributes of residential locations feasible. Simultaneously, through the integration of these physical characteristics with socio-economic indicators, a comprehensive understanding of the spatial distribution and heterogeneity inherent to these various settlement types can be attained. The approach employed in our study, which combines conventional data sources (such as census data and household surveys) with innovative spatial data derived from Earth observation, can be readily extrapolated to numerous other cities in the global South experiencing rapid urbanization within a multifaceted socio-economic milieu, yet confronted with a dearth of spatial data at the block or neighbourhood level.

In response to our research question i.e., does spatial inequity in access to services exist in Delhi? Our findings suggest that while spatial inequity in accessibility to services exists, the pattern of spatial inequity in Delhi differs from many cities in the global north. Unlike cities in the West, and especially in North America where suburban areas are home to high-income households and enjoy higher accessibility to many services, in Delhi we find the suburban neighbourhoods generally have poor accessibility to services. Also, a tight separation between urban and suburban areas in terms of high and low accessibility is not completely visible which makes the urban-suburban contrast in terms of services distribution weaker in cities like Delhi. Findings from other studies from the global south such as Rabiei-Dastjerdi et al. (2016) in their study of Tehran in Iran and Chen & Yeh (2019) in their study of Guangzhou in China also support this argument of a more dispersed pattern of spatial inequality in accessibility to services.

Answering our second research question, does spatial inequity exist on grounds of the neighbourhood's spatial location or socio-economic status? The result shows that while the neighbourhood's socio-economic characteristics do influence its accessibility to services, the neighbourhood's spatial location has a much larger influence as evidenced by the high magnitude of the DDC coefficient in the GWR model. The presence of socio-economic inequity inaccessibility is a common finding that exists in studies both from the global North and the global South. Many previous studies such as Zhao et al. (2020) in their study of Beijing, Saroj et al. (2020) in their study of different metro cities in India, Cortes (2021) in its study of Santiago in Chile, Bittencourt & Giannotti (2021) in their study of three cities of Sao Paulo, London and New York City, and Jin & Paulsen (2017) in their study of Chicago Metropolitan area find that households from low income and socially marginalised groups experience low accessibility to services.

While agreeing with findings from the literature, our study results show that the influence of socio-economic characteristics on accessibility is location-specific and cannot be generalised for the entire study region, due to the non-stationary relationship between the variables. We find that low-income neighbourhoods tend to have low accessibility to jobs and entertainment facilities, but in some locations, such neighbourhoods also have high accessibility to schools and hospitals. Similarly, neighbourhoods with a high percentage of scheduled caste population tend to have low accessibility to schools and hospitals, but at some locations, such neighbourhoods also have high accessibility to jobs and entertainment facilities.

The use of a spatial model like the GWR model, in the analysis of inequity in accessibility allows for a more granular understanding of how the relationships change across the city which is not possible to detect with non-spatial models like OLS or inequity measures such as Gini index. As our study results show that both neighbourhoods' spatial location and socio-economic characteristics can influence the accessibility to services, it requires that the services planning should target first the neighbourhoods dominated by low-income households and socially marginalised groups to have socio-economic equity in access to services. With more work in this direction, we believe findings from this study can aid in the formulation of policy guidelines on spatial planning of public services in Delhi which currently does not exist.

The history of city planning in numerous urban areas is marked by an enduring conflict between economic growth and inclusivity. Scarce resources have traditionally been distributed in ways that primarily benefit the affluent, with the belief that as the city expands, it will eventually uplift the less privileged. Empirical evidence from various cities illustrates that this economic-centric approach to service planning, while fostering urban growth, has also led to a range of issues including residential segregation, gentrification, soaring land prices, and unplanned colonies.

To foster cities that not only fuel economic growth but also promote inclusive development, urban sociologists have centred their attention on the theory of justice in facility planning. Among the diverse justice theories, the equity school of thought stands out as one of the most widely adopted approaches for addressing spatial and social inequalities. At its core, the concept of justice from the equity perspective underscores the "demand for equity," with a strong emphasis on "public welfare" and the right to "equal opportunities" for various social groups when it comes to vital resources, goods, and the government's commitment to upholding these rights (Davidoff, 1965; Davidoff et al., 1970).

Another branch of justice theory is the critical-spatial school, which conceptualizes spatial justice as an integral component of social justice, establishing a socio-spatial dialectic that calls for equitable distribution of physical resources and associated services among urban residents (Dikeç, 2009). Spatial justice posits that justice and space are mutually influential. Injustices stemming from economic, political, and social factors manifest spatially, potentially resulting in an inefficient spatial structure. This school of thought builds upon civil rights and democratic theories and draws on social concepts like Lefebvre's notions of the "production of space" and the "right to the city." According to Soja's exploration of spatial theory, all human spatial forms are socially constructed, leading to spatially uneven outcomes.

Consequently, social processes give rise to divergent geographies, shaping environments that reflect disparities in wealth and power.

Our findings show that the distribution of services in Delhi does not adhere to the norms of spatial and social justice which make it necessary that urban planning be combined with the concept of equity. Currently, the facility planning in Delhi follows a macro approach whereas land is allocated to different services based on the zone's population and availability of vacant land. Incorporating the ideals of equity planning and spatial justice, the zonal plans should demarcate the neighbourhoods according to their socio-economic characteristics and not just demographic features in the planning of services.

4.7 Conclusion

The present study aimed to investigate the spatial disparities in accessibility to services and their relationship with neighbourhood characteristics in Delhi. To achieve this goal, accessibility to four types of services for all 4145 residential locations in Delhi was measured using the E-2FCA method. The neighbourhoods were characterised based on their proximity to the city centre and district centres, population density, percentage of scheduled caste population, and neighbourhood richness index. The latter was calculated based on the neighbourhood's built-up form, land price, and mean household income obtained from a sampled household survey. To examine the variation in accessibility with neighbourhood characteristics, both OLS and GWR models were employed. While both models indicated that the relationship between accessibility and neighbourhood characteristics is statistically significant, the GWR model performed better, capturing the variables' non-stationary impact on accessibility.

The present study yields valuable insights into the role of accessibility in achieving sustainable urbanisation in Delhi. The high spatial variation in accessibility to the four services analysed underscores the existence of spatial inequity in accessibility, which has a significant impact on people's commuting behaviour and residential choices. Furthermore, our study reveals a significant association between neighbourhood socio-economic characteristics and accessibility. The influence of neighbourhood economic status on accessibility indicates the presence of vertical inequity, particularly concerning accessibility to jobs and entertainment facilities. This inequity in accessibility based on income reduces economic opportunities for the poor and reinforces income inequality. Additionally, our study highlights poor accessibility for neighbourhoods with a high scheduled caste population in certain areas. This lack of accessibility for specific social groups can impede development opportunities and further deepen social imbalances. Thus, accessibility has a pervasive impact on a region's growth, including its socio-economic, and individual well-being.

This study makes a significant contribution to the literature on accessibility-linked urban studies, both in terms of methodology and case study. While numerous studies have analysed variation in accessibility at the neighbourhood level for cities in the Western world, few have done so for cities in developing countries, mainly due to a lack of spatial data at the neighbourhood level. Different metrics for calculating accessibility exist in the literature, but many fail to capture the demand potential. Finally, the use of a spatial regression model for regression modelling revealed that spatial autocorrelation in the variables is an important aspect that should be considered when dealing with spatial data to obtain unbiased regression coefficients.

Two limitations of the study should be noted. First, As the study has relied on a sample of household income data collected through a survey to construct the four categories of economic status, some limitations of the survey design do affect the robustness of our results. This includes limited survey sample size, human bias in reporting the household income, error in data filing, and inaccuracy in survey location sites classification and stratified sampling procedure. Despite the survey design limitations, the constructed neighbourhood richness index stands as a good alternative approach to measuring neighbourhood economic status. Second, neighbourhood characteristics related to household composition, demographic structure, and racial/ethnic groups were not incorporated in the model, which may provide a better understanding of the impact of neighbourhood social composition on accessibility. With the availability of more detailed socio-economic data at the neighbourhood level in Delhi, the study can be further refined.

Chapter 5

Exploring Residential Built-Up Form Typologies in Delhi: A Grid-Based Clustering Approach towards Sustainable Urbanisation

Published as: Marwal, A., & Silva, E. A. (2023). Exploring residential built-up form typologies in Delhi: a grid-based clustering approach towards sustainable urbanisation. npj Urban Sustainability, 3(1), 40. https://doi.org/10.1038/s42949-023-00112-1

Chapter Overview: Urbanisation in developing countries presents opportunities and challenges for sustainable living beyond its relationship with economic growth. Studies in urban morphology demonstrate a significant linkage between urban form and sustainability, which can be leveraged to achieve sustainable urbanisation. However, the study of urban form in developing cities is limited in literature due to the diversity of micro-scale urban form features unique to these areas which makes the applicability of western urban form typologies challenging. Furthermore, the absence of spatial data and maps at the neighbourhood level presents a challenge in exploring micro urban form features. This study aims to uncover different residential built-up form typologies in Delhi using a grid-based k-means clustering algorithm and evaluate their impact on sustainable urbanisation. The clustering algorithm measures and visualizes variations in urban form by dividing residential areas into 100x100 meter grid cells and assigning attributes of accessibility, built-up density, and street design. The results reveal the presence of six unique built-up form typologies in Delhi. Based on the considered sustainable urban form elements, the study finds that only 19% of residential areas in Delhi can be classified under sustainable urbanisation while the rest of the areas require different sorts of planning interventions. The study methodology can be generalised to assess the micro-scale urban form features in cities of the global south that can provide a novel perspective to study urbanisation. Subsequent enhancements can be achieved through the diversification of urban form elements, including socio-economic components.

5.1 Introduction

Cities have significantly impacted human and environmental well-being throughout modern civilisation. Currently, more than half of the global population, which is 4.4 billion people, resides in urban areas, and this is projected to rise to 68% by 2050. According to the UN DESA (2018) report, Delhi is currently the second most populous urban agglomeration (UA) globally and is anticipated to surpass all other UAs in terms of population with an estimated 37.2 million individuals by 2028.

As the world becomes more urban, human interaction with a city's built-up environment is bound to increase, and thus important to study. The built-up of a city referred to as the urban form, provides an objective tool to understand this human-city relationship (Abrantes et al., 2019). Previous studies have emphasized the significance of comprehending urban form as a fundamental component of urban sustainability (Sharifi, 2019; Khavarian-Garmsir et al., 2023). Research shows that urban form influences a city's land use pattern and has a widespread impact on resident's lifestyle choices and the urban environment, such as residential location and commuting (Engelfriet & Koomen, 2018), social well-being (Mouratidis, 2018), environmental well-being (Zhou et al., 2018; Hankey & Marshall, 2017), and energy use (Zhang et al., 2019). Thus, a more comprehensive understanding of urbanisation can be achieved by characterising it through urban form, which is today largely driven by population density and the nature of employment (Eurostat, 2018).

While urban morphology has been a subject of academic inquiry for a considerable period, it has regained emphasis since the 1990s with the advancement of geographical information systems (GIS) and remote sensing (Abrantes et al., 2019). Over the last 20 years, studies have used different classification methods to categorise the morphological elements of urban form at different spatial scales such as neighbourhoods and cities, applying both quantitative and qualitative categorisation tools (Fleischmann et al., 2021). While these studies enrich our understanding of the relationships between different urban configurations and their impact on urban sustainability, the bulk of our comprehension regarding urban morphology has come from the cities in the global north. However, in recent years a discernible shift towards understanding urban morphology in cities in the global south has been noticed. (Peimani & Kamalipour, 2022; Xu et al., 2019).

In recent decades, cities in developing countries have seen a very high influx of migrants, leading to changes in land use patterns and the mushrooming of diverse residential settlements, many of which are often in unregularised neighbourhoods and can be termed illegal settlements or slums (Sandoval & Sarmiento, 2020). In cities with such diverse settlement patterns, neighbourhoods differ not only in terms of their socio-economic indicators but also in their built-up structure (Kraff et al., 2020). In such cases, characterising a city with a particular urban form can be misleading. Thus, to study the extent of sustainable urbanisation in such cities, one needs to explore the variation in neighbourhood built-up types. Although a topic of great importance, as Mahtta et al. (2019) point out in their analysis of urban forms of 478 cities, few studies have analysed variations in urban form within a city.

In this context, this study raises some important questions - Do residential areas in Delhi have diverse built-up forms? If so, how can we visualise and measure them? Moreover, what impact does the built-up form have on sustainable urbanisation in Delhi? In this regard, the study has two objectives, first, to cluster the neighbourhoods in Delhi using the k-means clustering algorithm and characterise them with their dominant built-up form and typology. Second, to analyse how these different built-up form typologies affect sustainable urbanisation. It is important to note that as the study aims to explore residential built-up form typologies, it considers only the physical aspects of residential areas. Other aspects of residential areas related to socio-economic and demographic features have not been explicitly addressed in this study.

In the context of urban morphology, to the best of our knowledge, this study is one of the earliest to examine and map the variations in the built-up form of residential areas in Delhi. The study holds significance on two grounds. First, the study uses the grid-based clustering method which provides a methodological tool for urban planners to delineate the different built-up form in the city in a more dynamic and adaptive manner in contrast to relying on administratively defined boundaries. This provides more flexibility in mapping as it can easily adapt to changes in the city's physical layout and demographics. The method also offers greater precision as it considers the actual physical layout of the city in the mapping (Leasure et al., 2020). With flexibility and precision, this method also offers consistency in mapping residential areas across different cities and regions which makes it generalisable and thus, significant for cities in the global south that lack micro-scale spatial maps.

Second, and more importantly, by addressing the sustainable cities paradigm, as specified under sustainable development goal 11: sustainable cities and communities, the study provides a novel perspective to study urbanisation through the lens of urban form. While urbanisation is seen as a synonym for economic growth, if poorly planned it can have adverse implications for individual and environmental well-being (Nieuwenhuijsen et al., 2020). In this context, we argue that the study of urbanisation should include elements of urban morphology (Schirmer & Axhausen, 2019). By doing so, we can have a more informed understanding of how future urbanisation will impact the neighbourhood’s living and what planning interventions can be made to achieve sustainable urbanisation. This can help ensure that urbanisation leads to sustainable cities rather than just economic growth.

The remainder of the paper is organised as follows: Section 5.2 provides a comprehensive literature review that examines existing issues in analysing urban forms. In Section 5.3, the data preparation and research methodology are thoroughly explained. The study's findings are presented in Section 5.4. Section 5.5 discusses the study results and their implications for promoting sustainable urbanisation in Delhi. Section 5.6 concludes the paper.

5.2 Literature Review

Our examination of existing literature on urban form reveals three primary issues. First, studies conducting systematic exploration to capture the heterogeneity of spatial patterns at the neighbourhood level are found to be limited (Fleischmann et al., 2022). While previous studies have analysed cities based on their dominant urban forms (Su et al., 2021; Li et al., 2021), few have applied quantitative methods to study the variation in urban forms within a city (Mahtta et al., 2019; Masoumi et al., 2019). Cities, especially in developing countries, have diverse settlement patterns that result from inadequate zoning laws and weak regulations (Debray et al., 2023; Ahmad & Choi, 2011). Unplanned urbanisation can result in the proliferation of urban sprawl, slums, and unauthorised colonies in a city, which has distinct urban forms compared to more affluent areas (Abascal et al., 2022). Thus, it is crucial to understand the possible urban form typologies within a city for effective localised land use planning (Tang et al., 2021; Palaiologou et al., 2021).

A few recent studies have investigated urban form typologies at the neighbourhood level. For example, Braulio et al. (2018) developed a taxonomy of the city of Castellón de la Plana, Spain, using elements of residential buildings and analysed variations in urban form patterns at different geographical scales. Lu et al. (2019) measured urban form in different neighbourhoods of Chengdu City, China using indicators of density, accessibility, shape and diversity. Fleischmann et al. (2022) developed a numerical taxonomy for urban form to classify urban types using street networks and building footprints which they applied to generate a hierarchical classification of urban form in Parague and Amsterdam. On similar lines, Fusco et al. (2022) built a taxonomy of contemporary urban forms in France using indicators of street design and building types. With limited studies analysing urban form in the global south, our study contributes to the growing body of literature on the built-up form typologies at the neighbourhood level in cities in the global south.

The second issue concerns the geographical scale for measuring urban form an inappropriate selection of which can lead to the modified area unit problem (MAUP). The Modified Area Unit Problem (MAUP) refers to the phenomenon wherein the results of statistical analysis vary based on the scale or size of the geographic units used (Openshaw, 1984). It is a common issue in spatial analysis where geographic data can be aggregated into different levels of spatial resolution, such as census tracts, counties, states, or countries. Studies find, based on the choice of spatial resolution results can vary (Viegas et al., 2009).

To minimise the MAUP, data should be aggregated at the most appropriate spatial scale as per the research objective. In urban micro-planning, residential blocks/ neighbourhoods can serve as a unit of spatial analysis. However, if block-level spatial maps are unavailable, a grid cell approach can be used to define the study area (Hou, 2016; Nedovic-Budic, 2016). Under this approach, the geographic space is divided into a mesh of identically sized cells that are commonly square-shaped, known as a grid. Each cell contains a numerical value that represents a specific geographic attribute, such as density or elevation, for that unit of space (Tollefsen et al., 2012).

One of the key advantages of the grid cell approach is that it allows for a high degree of precision and granularity in spatial analysis (Leasure et al., 2020). By breaking down a geographic area into small cells, researchers can identify even subtle spatial patterns and relationships that might be missed with other methods (Brown et al., 2019). Aggregating data under grid cells has been found to yield better results than aggregating data under administrative boundaries (Rothlisberger, 2017; Ahmed & Bramley, 2015). Considering the lack of administratively defined neighbourhood boundary maps in Delhi and the advantages of the grid cell approach in minimising MAUP, this study relies on the grid cell approach as a spatial unit of analysis.

Finally, there is a challenge in clustering spatial data in urban form analysis. Clustering techniques in spatial science refer to a group of methods used to identify and group spatially related data points or objects. Clustering involves grouping objects based on their similarity in terms of geographic attributes, such as distance, spatial density, or other spatially relevant features (Jiawei et al., 2012). The basic idea behind clustering techniques is to divide a dataset into subsets or clusters such that the objects within each cluster are more similar to each other than to objects in other clusters. This allows researchers to identify spatial patterns or groupings in the data that may not be immediately apparent through visual inspection. Apart from analysing the cities' physical characteristics, cluster analysis has also been used to characterise cities based on their planning objectives (Xu & Heikkila, 2020), and socio-economic and demographic features (Cheng et al., 2021).

Recent studies on urban morphology have applied various clustering techniques to group spatial units with similar attributes and to identify dominant urban form typologies. These techniques include k-means, hierarchical agglomerative clustering (Su et al., 2021; Mehrotra, 2019; Oke et al., 2019), density-based clustering (Pilehforooshha & Karimi, 2019), Bayesian clustering (Guyot et al., 2021), Gaussian mixture model (Jochem et al, 2020), and spatial clustering methods like local indicators of spatial association (LISA) and local indicators of network-constrained clusters (ILINCS) (Perez et al., 2020). Advanced methods like self-organising maps, which combine statistical and machine learning methods, are also being used (Abrantes et al., 2019; Li et al., 2019).

The study after using different clustering algorithms, realised k-means clustering to be more suitable in terms of model construction and execution. The model uses few input parameters and can be executed using different open-source software (Yuan & Yang, 2019; Fränti & Sieranoja, 2019). The model results are also easy to interpret compared with hierarchical clustering (Govender & Sivakumar, 2020). We also noticed the model requires less computational time and is more efficient in processing large multivariate datasets in comparison to density based and agglomerative clustering (Patel & Kushwaha, 2020; Ahmed et al., 2020). More importantly, we found that among different clustering techniques, k-means clustering has been widely used in recent studies (Iqbal et al., 2022; Shi et al., 2021; Xu et al., 2021; Bobkova et al., 2021; Schirmer & Axhausen, 2019). Thus, owing to its widespread applicability, it was easier for us to connect our model results with studies that have employed similar clustering techniques.

However, one of the limitations of the K-means algorithm is that it requires the number of clusters to be specified in advance, which can be difficult if the data does not have a clear structure (Patel & Kushwaha, 2020). Incorrectly specifying the number of clusters in k-means clustering can produce oversimplified or overcomplicated results. A low number of clusters may miss important distinctions between data points and result in the loss of information, whereas an excessive number of clusters may lead to meaningless clusters, obscuring the underlying structure of the data (Fahim, 2022). To overcome this limitation, this study uses a cluster optimisation method which is discussed in the methodology section.

This section concludes by highlighting the challenge of interpreting the cluster results. Although machine learning (ML) based clustering algorithms can cluster multivariable big datasets, they may not provide a clear understanding of how to interpret the cluster results (Murdoch et al., 2019). With the increase in the usage of machine learning models in data analysis, there has been growing concern about how to efficiently interpret the results of ML models (Rai, 2020). This becomes particularly challenging when multiple features are significant predictors of a cluster, making it difficult to determine how a particular feature impacts the cluster prediction (Brandsæter & Glad, 2022).

The ML-based clustering algorithms typically employ unsupervised learning techniques, meaning that there is no predetermined outcome. As a result, the interpretation of the results is subjective and relies heavily on the analyst's understanding of the data and research question. Thus, unfamiliarity with the working of ML models or inadequate understanding of the study context can lead users to wrongly interpret the model results which can affect the study findings (Wang & Biljecki, 2022).

To overcome this issue, a recent development has been the rise of explanatory methods, such as SHAP to enhance the interpretability and transparency of ML models (Khadem et al., 2022). SHAP (SHAPley Additive exPlanations) is a method that explains the output of a machine learning model by computing the contribution of each feature to the final prediction by averaging the marginal contribution of each feature over all possible coalitions of features (Lundberg & Lee, 2017). The SHAP method is effective in providing insights into the inner workings of complex machine-learning models (Ekanayake et al., 2022). In the last few years, the use of the SHAP tool can be seen in studies from different disciplines such as health (Zheng et al., 2021), engineering (Meddage et al., 2022) and finance (Mokhtari et al., 2019) which shows the growing acceptance of the SHAP tool in interpreting cluster results.

5.3 Methodology

5.3.1 Study context

Delhi, the capital city of India, is the 2nd most populous city globally, with over 28 million residents. It is expected to become the densest city in the world by 2030 (UN DESA, 2018). Spread over 1483 sq. km, Delhi is divided into 11 districts and 250 wards that come under the Municipal Corporation of Delhi (MCD), with almost 3000 residential colonies. Figure 5.1 shows the map of Delhi with the residential areas under the 100x100 meter grid size. The city is a hub for social, economic, and cultural activities, attracting migrants from across the country and abroad. In the last decade, Delhi's built-up area has almost doubled, making housing and transportation management challenging for the government (Naikoo et al., 2020). According to Delhi's 2018-19 socio-economic survey, 85% of the population requires affordable housing options, while 11% live in slums, and 60% of households (size of 5) are congested.

The 2011 Census estimates that Delhi will need 34.5 lakh dwelling units by 2041. As per the Delhi Development Authority (DDA, 2022), the mismatch between housing demand and supply and unaffordable prices has led to the growth of over 1700 unauthorised colonies in the city, housing over 4 million people. The Delhi economic survey 2019-20 reports that there are 643 vehicles per 1,000 population, double the 2005-06 number, leading to traffic congestion, road accidents, and parking space shortages. The Master Plan Delhi-2041 projects that Delhi will have over 46 million daily trips, with a per capita trip rate of 1.58.

These are some of the issues which Delhi is facing due to massive urbanisation, which may be exemplified further if suitable planning interventions are not made. In this context, we take Delhi as our case study to understand what local planning interventions can be done to achieve sustainable urbanisation.

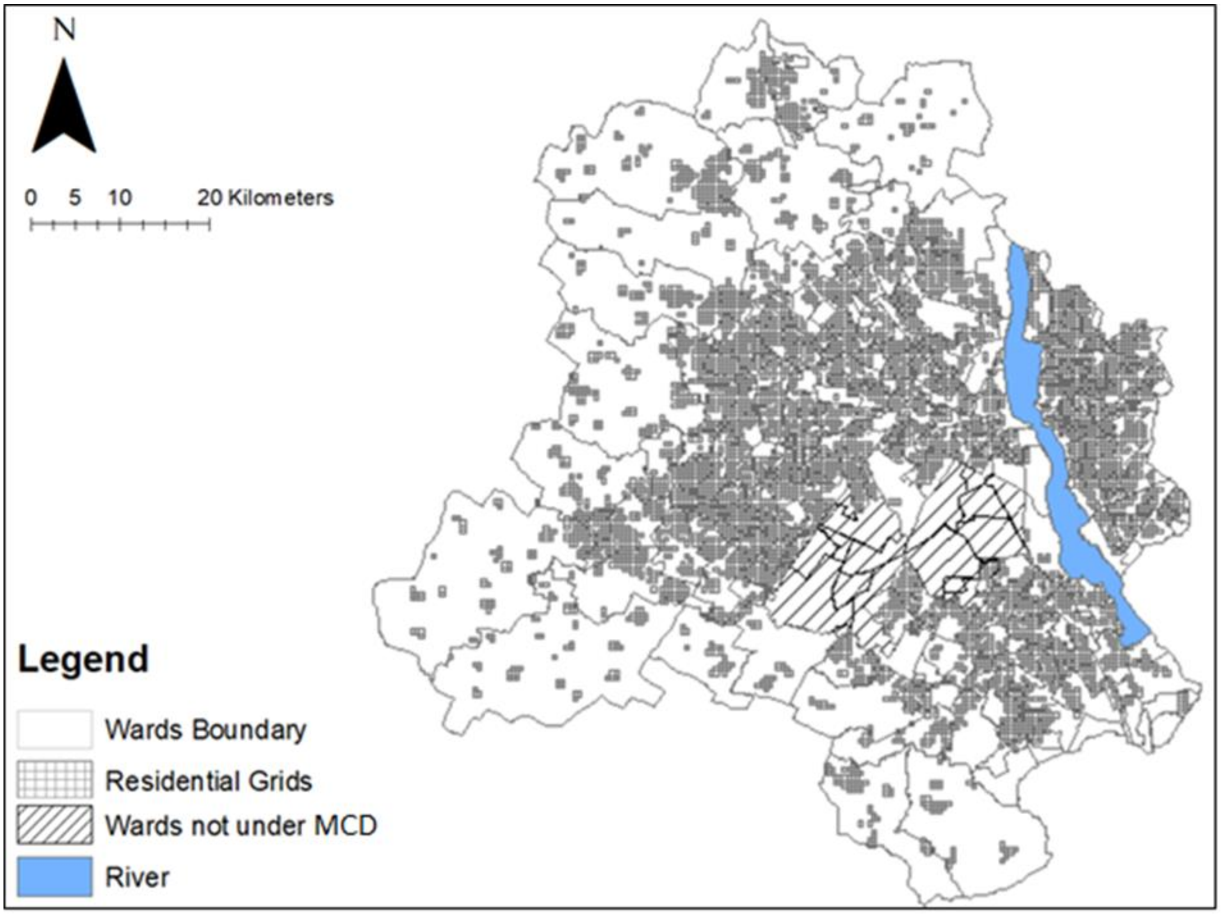

Figure 5.1: Map of Delhi with residential areas shown under grids.

5.3.2 Data preparation

The first step in the data preparation process involved mapping residential areas in Delhi using Google Earth. The mapped residential area was then exported to ArcGIS software and transformed into a raster file with a cell size of 100 x 100 metres, resulting in 37,092 grid cells. These grid cells served as the spatial unit of analysis and were assigned different urban form elements.

Urban form refers to the physical and spatial characteristics of urban areas, including the arrangement and distribution of buildings, streets, open spaces, and other features that shape the built environment (Anderson et al., 1996). Elements of the urban form commonly found in previous studies belong to the

5D framework developed by Ewing & Cervero (2010), which includes, population density, land use diversity and street design, destination accessibility and distance to transit stations (Bourdic et al. 2012; Sharifi, 2019). Despite its widespread use, there has yet to be a consensus in the literature on what constitutes urban form. As Fleischmann et al. (2021) note, the term has many interpretations, leading to the need for an objective system of measuring urban form features. In this study, we focus on commonly used elements: (a) density, including built-up density and growth in built-up density from 2012 to 2022, (b) street design, including street intersection density and block size, and (c) accessibility to five services. Our choice of elements is based on the research aim, study area context, data availability, and ease of interpretation.

The first element is the accessibility to services. While accessibility can be computed using factors like time, distance, or demand-supply, we stick to the cumulative measure of accessibility that is easy to measure and interpret and provides a good indication of the spatial distribution of services in a neighbourhood (Kelobonye et al., 2019b). A buffer radius of 2 km was used to represent services lying in the immediate neighbourhood that can be accessed with a non-motorised travel mode. We considered five types of services: schools, hospitals, entertainment facilities, commercial areas, and metro rail stations. Good access to schools and hospitals is crucial for human development, as highlighted in Sustainable Development Goals 3 and 4. High accessibility to metro stations provides quick and convenient transportation options to people, reduces their travel time, and improves overall mobility. Studies show that high access to entertainment facilities and commercial areas enhances the quality of life (Li et al., 2021).

Table 5.1 lists the services, the number of observations for each service, and their sources. To compute the accessibility to each of these services, we first obtained the location addresses of all the observations from their respective sources and then created a spatial database by geocoding the addresses in Google Maps. The geocoded addresses of these observations for every service were then mapped in ArcGIS in a point shapefile format. Grid cells were assigned a total count of observations lying in a circular radius of 2 km for every service. Finally, accessibility was calculated as the sum of the normalised value of the cumulative count of observations for all five services, as shown in Equation (5.1).

$$A_i = K \sum_{j=1}^{5} \left[\frac{x_{ij} - min_j}{max_j - min_j} \right] \tag{5.1}$$

Where A_i represents the accessibility of the grid cell i, x_{ij} represents the total count of observations belonging to service j and lying in the circular radius of 2 km from the grid cell i, min_j and max_j represents the minimum and maximum count of observations, respectively, belonging to service j and lying within a radius of 2 km across the grid cells. K was used as a constant to keep the accessibility values within a reasonable limit.

Table 5.1: Summary of services data and source.

Services	Number of Observations	Variables	Source
Schools	2403	Location of private and government schools of all levels.	Directorate of Education, Government of Delhi.
Hospitals	969	Location of public and private hospitals offering tertiary care	Department of Health and Family Welfares, Govt. of Delhi
Entertainment Facilities	827	Location of shopping malls, movie theatres, and registered restaurants	Department of Excise, Entertainment & Luxury Tax, Govt. of Delhi
Metro Stations	185	Location of all metro stations	Delhi Metro Rail Corporation
Commercial areas	23	District centres and Sub-district centres	Delhi Master Plan 2020

The second and third factors were built-up density and growth in built-up density, respectively. The built-up density was calculated in three sequential steps. (a) land satellite imagery of Delhi for the year 2022 was acquired from the USGS (United States Geological Survey) and exported in ArcGIS. (b) Land use classification was done using a supervised classification tool to classify land use under built-up, bare soil, cultivated area, wasteland, and water. The accuracy of land use classification was verified by cross-checking the land use of randomly sampled 350 data points from the classified image with the actual land use as visible in the historical imagery tool of the Google Earth software. The classification accuracy measured using the kappa statistics was 0.86, which denotes a high accuracy of classification (Girma et al., 2022). (c) The area under the built-up category was extracted from the classified image and vectorised into points. The built-up density of a grid cell was calculated as the number of built-up points lying in a 500 metre buffer radius from each grid cell.

The same process was repeated to calculate the built-up density for each grid cell in 2012. Finally, we calculated the growth in built-up density as the percentage change in built-up density from 2012 to 2022 for each grid cell.

The fourth and fifth elements of our study were street intersection density and block size, respectively, which were computed using the open street map (OSM) database. The OSM is a collaborative open-source mapping platform that provides information on roads, buildings, landmarks, and other geographic features (OSM, 2023). OSM can be accessed through various applications, such as QGIS, which was used in this study. To calculate street intersection density, we used the residential street layout from the OSM database, which provides a map of all residential streets in Delhi. Using the line intersection tool in QGIS, we measured the number of residential street intersection nodes within a 500 metre radius for each grid cell.

To measure the block size, we first calculated the area of the polygons formed by enclosing residential streets. For a grid cell, the block size represented the average area of all polygons lying inside a buffer radius of 500 metre.

A data summary of all the elements of the urban form used in this study is shown in Table 5.2. The research methodology described in this section is summarised in the flowchart in Figure 5.2.

Table 5.2: Data summary

Variable	Mean	Std. dev.	Min	Max
Accessibility	134.8	115.3108	0	735.6301
Built-up area	245.5816	69.32736	15	348
Growth in built-up area	10.33771	35.44537	0	314
Street intersection density	116.738	91.47772	1	681
Block area (sq. meters)	5187.115	1682.68	1602.367	17379.24

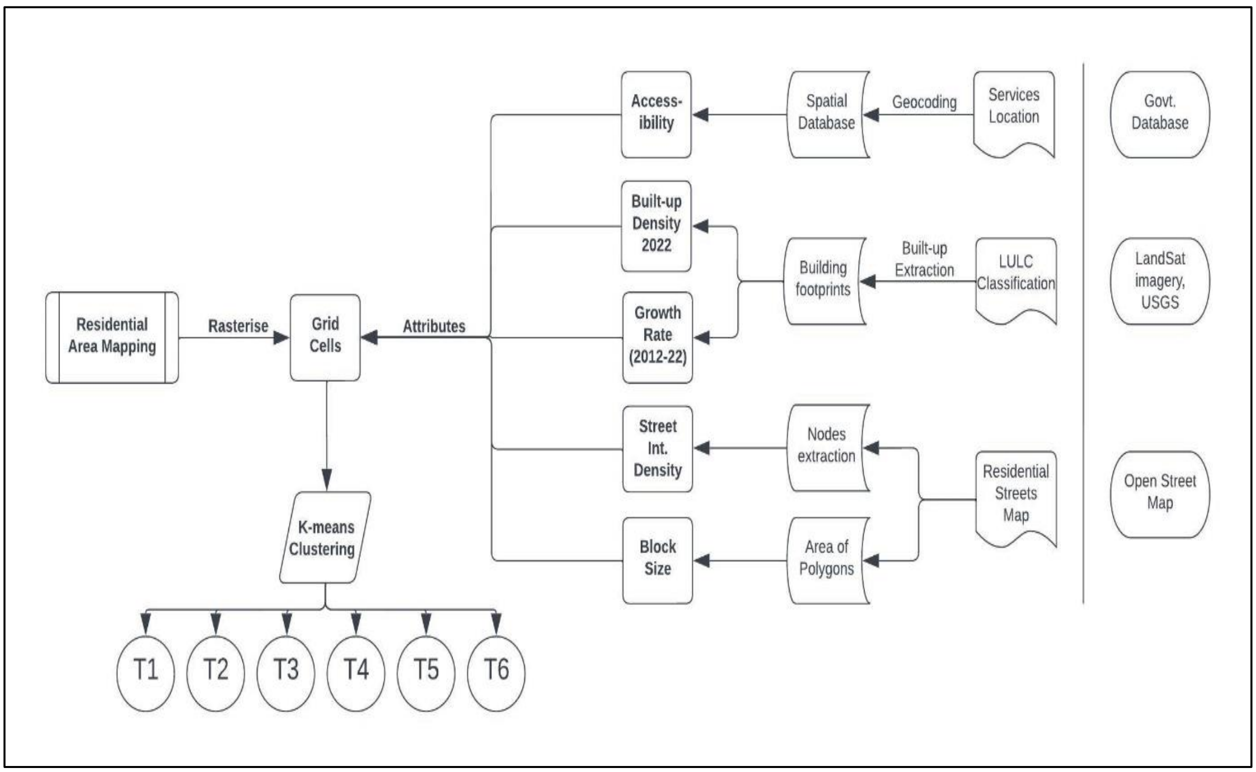

Figure 5.2: A flowchart of research methodology

5.3.3 K-means clustering

We used k-means clustering to classify the grid cells into similar urban form attributes. K-means clustering is a machine learning algorithm used for clustering or grouping data points in a dataset. The algorithm partitions the data into 'k' non-overlapping clusters, where k is a predefined number chosen by the user. The algorithm works by iteratively assigning each data point to the closest cluster centre (centroid) and then recalculating the centroid of each cluster based on the newly assigned points. This process continues until the centroids no longer move significantly, or a specified maximum number of iterations is reached (Jain et al., 1999).

We used the *scikit-learn* library (Pedregosa et al., 2011) in Python to execute the algorithm. First, we scaled the data using the min-max scaler to a range of 0 to 1. We then determined the optimum number of clusters using the widely used elbow method (Syakur et al., 2018). The elbow method calculates the total variation within a cluster using WCSS (within-cluster sum of squares) and plots the results to determine the optimum number of clusters (k).

As the number of clusters increases, the variation within every cluster, i.e., the value of WCSS, is expected to decline. The optimum value of k is one where the marginal decrease in the value of WCSS by adding one more cluster is minimal. This can be visualised by plotting the values of WCSS against the number of clusters (k). As shown in Figure 5.3, when the value of k is six, a sharp bend or elbow-shaped curve occurs in the graph. At this point, the marginal change in the value of WCSS with an increase in the value of k is minimum.

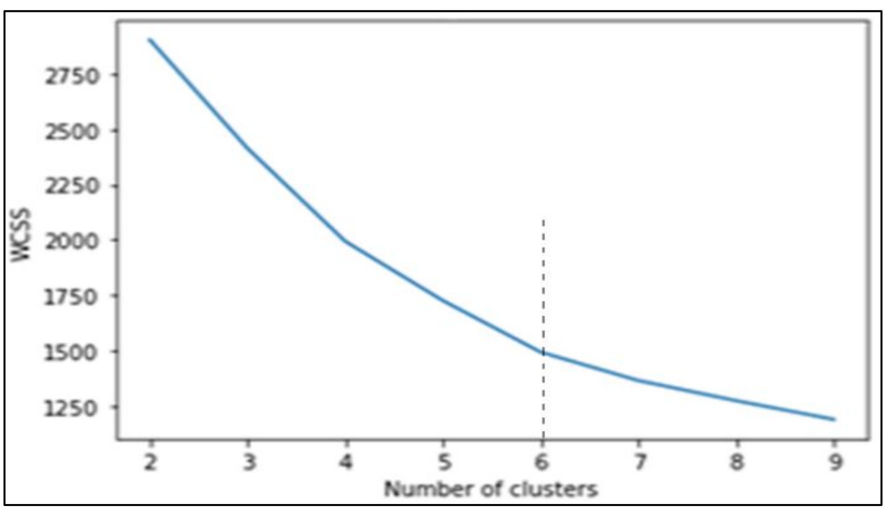

Figure 5.3: Elbow method to figure out the optimum number of clusters (the dotted line shows the optimum number of clusters)

The elbow method showed that the optimum number of clusters was six, which was used in the k-means clustering. Each cluster was then assigned a built-up form typology using the SHAP tool.

5.4 Results

5.4.1 Cluster characteristics

One of the study objectives was to cluster the neighbourhoods in Delhi and characterise them with their dominant built-up forms. After scaling the data and determining the optimum number of clusters, we ran the k-means clustering algorithm with five urban form elements. The algorithm clustered the grid cells into six distinctive clusters, designated T1 to T6. Table 5.3 shows the percentage share of the grid cells in the six clusters. Cluster T6 has the maximum share of the total grid cells, i.e., 28%, while cluster T3 has the least share, i.e., 7%.

Figure 5.4 shows the spatial distribution of the grid cells under the six clusters. We find cluster T1 is located in the centre and south of the city. Cluster T2 is spread in the south and north of the city, while cluster T3 is visible in small pockets in the north and west of the city. Cluster T4 is primarily located in the city's outer areas, and cluster T5 can be seen in locations lying towards the central west of the city. Lastly, cluster T6 is primarily clustered in the northeast of the city.

Table 5.3: Percentage share of grid cells in the six clusters

Cluster	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**
Share of grid cells (%)	19	23	7	10	13	28

To determine the statistical disparity among the clusters, we performed the multivariate analysis of variance and covariance (MANOVA) test. MANOVA is a statistical test used to measure the impact of one or more independent variables (factor variables) on two or more dependent variables. In other words, the MANOVA test determines whether the mean value of the dependent variable changes for different groups in the independent variable. The null hypothesis assumes that there is no statistical difference in the mean values of the chosen dependent variables across different groups (Johnson & Wichern, 2007).

We ran the MANOVA test using the urban form elements as dependent variables and the assigned clusters as independent variables. We used four different test parameters to determine statistical significance. All four tests computed in the MANOVA rejected the null hypothesis based on p-value significance. Table 5.4 presents the result of the MANOVA test. The test revealed substantial disparities in the average values of urban form elements among the six clusters, indicating that each cluster possesses a different composition of urban form features.

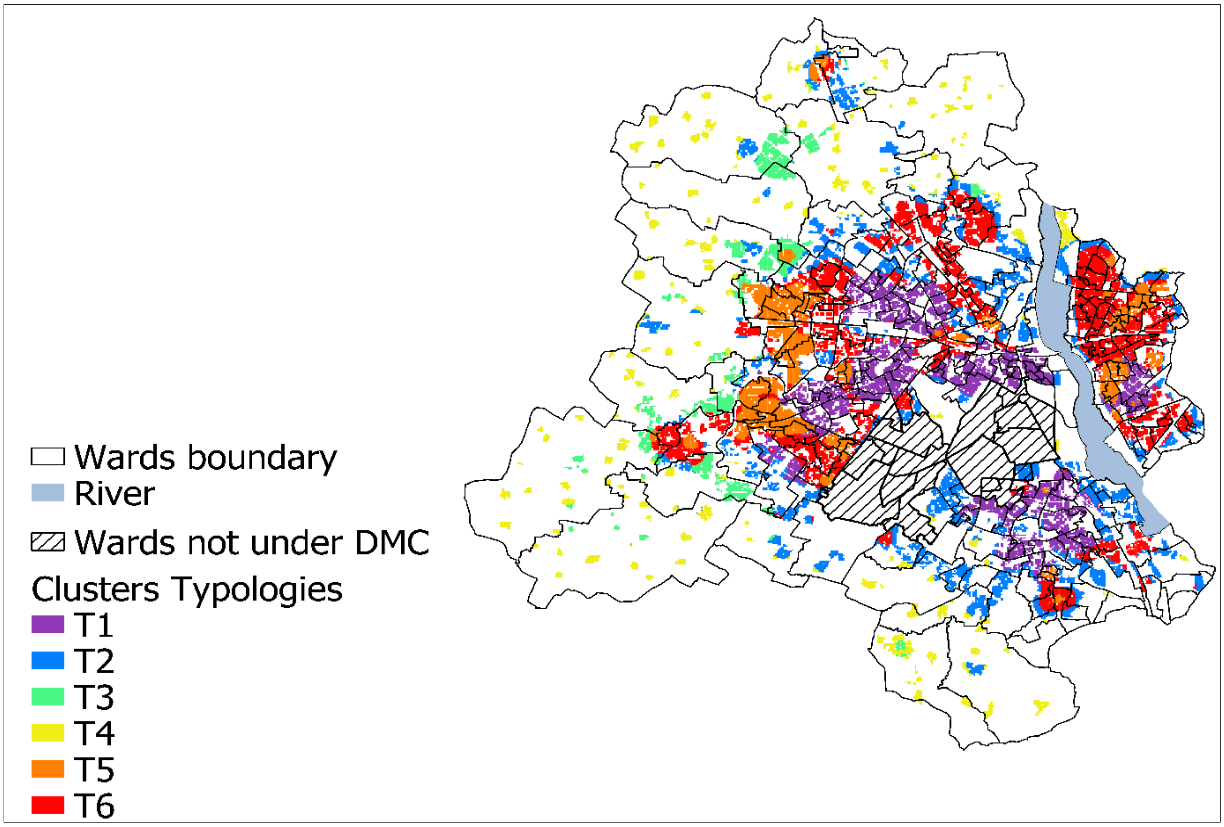

Figure 5.4: Spatial distribution of grid cells under the six clusters

Table 5.4: MANOVA result

Number of observations = 37,092					
W = Wilks' lambda L = Lawley-Hotelling trace					
P = Pillai's trace R = Roy's largest root					
Source	**Statistic**	**df**	**F(df1, df2) = F**	**Prob>F**	
Clusters	W 0.0145	5	25.0 137755.0 11707.02	0.00	a
	P 2.4759		25.0 185430.0 7275.75	0.00	a
	L 8.8455		25.0 185402.0 13119.79	0.00	a
	R 4.8066		5.0 37086.0 35651.56	0.00	u
Residuals	37086				
Total	37091				
e = exact, a = approximate, u = upper bound on F					

To graphically analyse the differences among the clusters, two data visualisation techniques have been employed, which include box plots and parallel coordinate plots. Figure 5.5 (a-e) displays the box plots of various urban form elements across the six clusters. In a box plot, the distribution of the data is represented using a box and a set of whiskers. The box in a box plot represents the interquartile range (IQR) of the data, which is the range between the 25th and 75th percentiles of the data. The median value of the data is represented by a line inside the box. The whiskers extend from the box to the minimum and maximum values of the data, excluding any outliers. Looking at Figure 5.5 (a-e) one can notice how the data distribution of variables differs for different clusters, marked from T1 to T6. For example, Figure 5.5 (a) shows median accessibility score is highest in cluster T1, while Figure 5.5 (b) shows median street intersection density is highest in cluster T5.

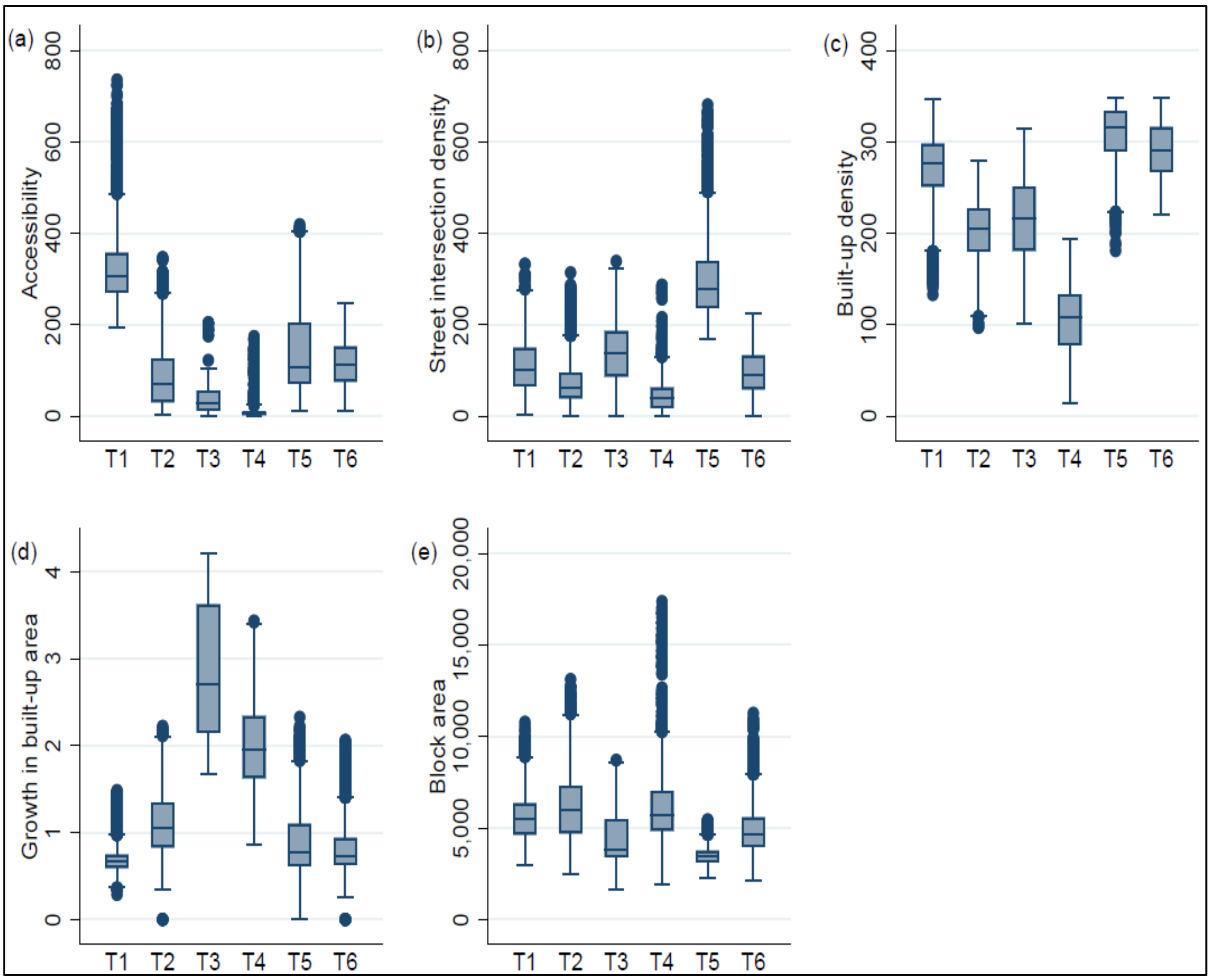

Figure 5.5**:** Box plot analysis for (a) accessibility, (b) Street intersection density, (c) Built-up density, (d) Growth in built-up area, and (d) Block area

The parallel coordinate plot graphically depicts the multivariate data of each cluster, offering a visual illustration of the disparities between the clusters. In a parallel plot, each variable in the dataset is represented by a separate axis, which is arranged in parallel to each other. The data points are then plotted as a set of connected line segments across the different axes, with each line segment representing the value of a particular variable for a specific data point (Moustafa et al., 2011). As can be seen in Figure 5.6, the urban form variables have been marked on the x-axis and the clusters are shown with different coloured line segments. For example, cluster T1 (marked in green) has the highest score for accessibility, while cluster T3 (marked in red) has the highest score for growth in the built-up area.

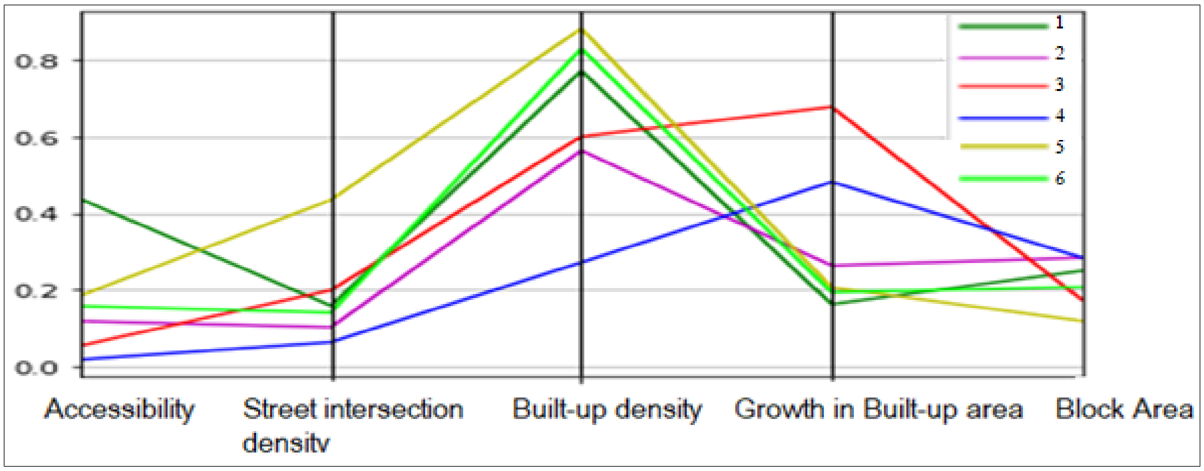

Figure 5.6**:** Parallel coordinate plot (the variables are shown on the x-axis, and their normalized mean value score for every cluster is shown on the y-axis).

5.4.2 Cluster typology

The SHAP (Shapley Additive exPlanations) tool was employed to visualise the variations between the clusters using the *shap* module in Python. As described in section 2, the SHAP tool is a machine learning technique that provides a way to explain the contribution of each feature in a prediction made by a model, and how that feature affects the output. The SHAP tool generates a plot called a "summary plot" that displays the most important features and how they affect the prediction. The summary plot ranks features based on their contribution to cluster characterisation, with the most important features at the top. Each feature is represented by a horizontal bar where the colour of the bar indicates the value of the variable or feature, with blue indicating a low value and red indicating a high value (Lundberg & Lee, 2017). The direction of the SHAP value denotes how effectively a feature value can characterise the cluster, where a positive SHAP value denotes a feature that is more likely to characterise the cluster. Figure 5.7 displays the summary plot of the SHAP tool.

The summary plot shows the nature and magnitude of the impact of the urban form elements on cluster prediction. The element that has the strongest impact on cluster predictability is considered as dominant element and used to frame the cluster typology. In the case of Cluster T1, we find that high values of accessibility to services, built-up density, and block size have a positive SHAP value. Moderate values of street intersection density and low values of the growth rate also have a positive SHAP value. This means that cluster T1 is more likely to be predicted by high values of accessibility to services, built-up density, and block size, along with moderate values of street intersection density, and low values of growth.

However, to construct the cluster typology, we use the cluster feature that has the strongest impact, which is a high value of accessibility to services in the case of cluster T1. Thus, we label cluster T1 as an area with high accessibility to services.

Based on the analysis of cluster T1, the typologies of the remaining clusters can be similarly framed. Cluster T2 has a moderate built-up density as its dominant feature, and it is also marked by a moderate growth rate, a low accessibility value, and a large block size. With such characteristics, cluster T2 is labelled as a moderate built-up density area. The dominant feature in cluster T3 is a very high growth rate, with a moderate built-up density and street intersection density also contributing to cluster characterisation. Thus, cluster T3 is labelled as an area with a high growth rate.

In cluster T4, low built-up density is the dominant feature. Other important features include a high growth rate and low accessibility. We find that residential areas in this cluster are mostly located on the city's periphery. Considering the high growth rate and spatial location of residential areas lying in cluster T4, the cluster is referred to as an urbanisable area with a rural landscape. The dominant feature in cluster T5 is high street intersection density. High built-up density and small block size also contribute to the cluster characterisation, which makes the cluster densely populated. Thus, cluster T5 is labelled as a compact and congested area. Finally, in cluster T6, we find that high built-up density has the highest positive impact, followed by moderate to low accessibility and moderate to low street intersection density. Due to this, cluster T6 is referred to as a high density area.

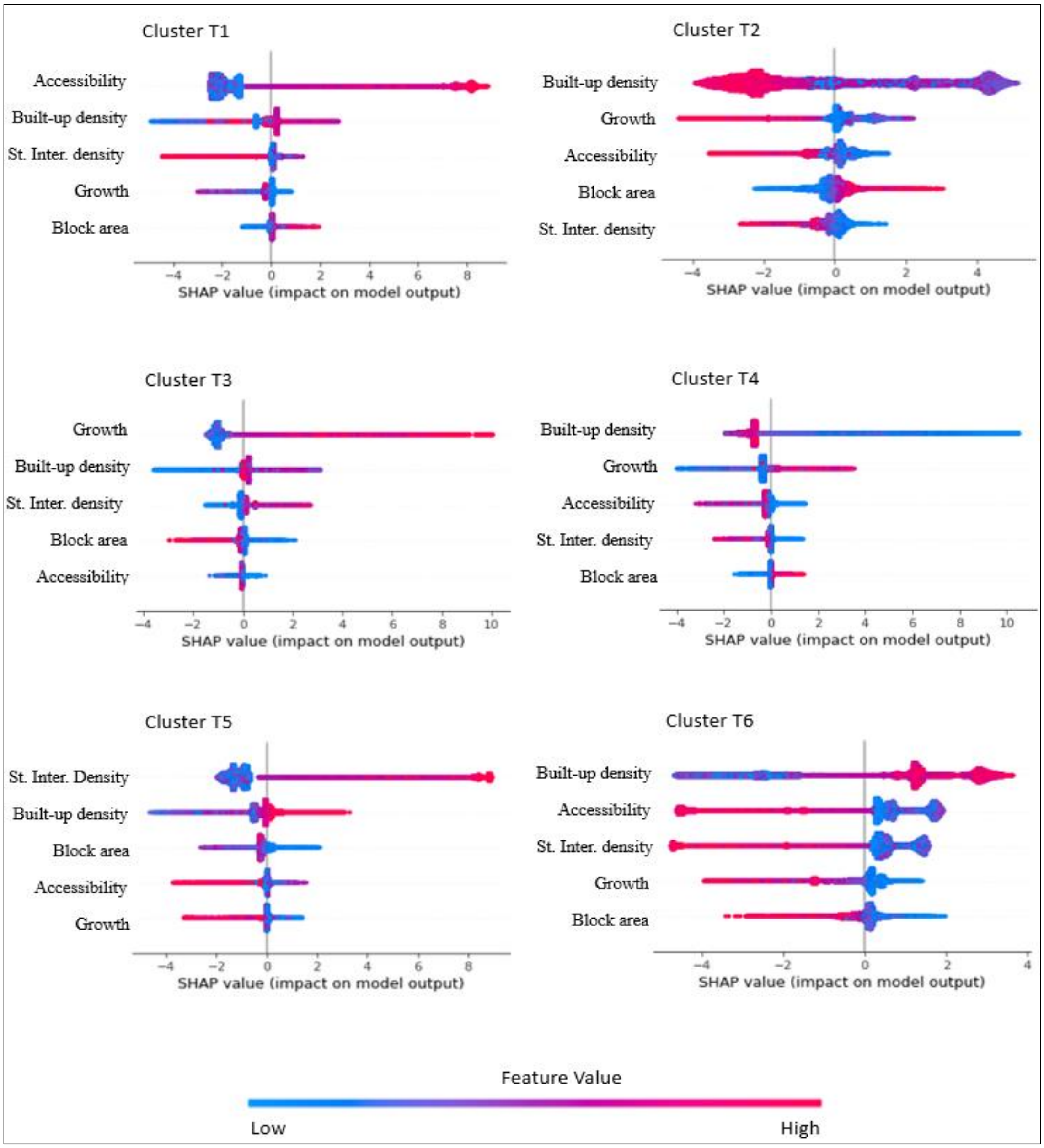

Figure 5.7: SHAP value plots

Table 5.5 tabulates the clusters' characteristics based on the boxplot and multivariate plot analyses, and their dominant urban form or typology framed using the SHAP tool. The next section discusses the clusters' characteristics and their impact on sustainable urbanisation.

Table 5.5: Cluster characterisation and typology

Cluster	**Accessibility**	**Built-up density**	**Growth rate**	**Street intersection density**	**Block Area**	**Dominant Urban Form or Cluster Typology**
T1	+	+	-	+/-	+	High accessibility to services
T2	-	+/-	+/-	-	+	Moderate built-up density
T3	-	+/-	++	+/-	+/-	High growth rate
T4	-	-	+	-	+	Urbanisable area with a rural landscape
T5	+/-	+	-	++	-	Compact and congested
T6	+/-	+	-	+/-	+/-	High density

Before concluding this section, we highlight a few recent studies that have used k-means clustering to analyse neighbourhood typologies. For example, Vogiazides & Mondani (2023) used k-means clustering to cluster neighbourhoods in Sweden to analyse variation in neighbourhood status and found ten different neighbourhood types. Wu et al. (2022) used k-means clustering to identify four neighbourhood typologies for London, Paris and Amsterdam. Similarly, Lynge et al. (2022) used k-means clustering to build eight neighbourhood typologies for different cities in South Africa. Although due to different study contexts and choice of variables the results from these studies cannot be directly compared with our study results, the fact that k-means clustering has been used successfully in these studies reinforces the reliability and robustness of this technique. The demonstration of consistent findings in prior studies lends additional support to the validity of the current research.

5.5 Discussion

The cluster analysis identified six distinct residential form typologies in Delhi. This section develops on the study's second objective, which is to analyse how these different built-up form typologies affect sustainable urbanisation in Delhi. Before discussing further, we first list the parameters used in this study to define sustainable urbanisation. Sustainability is a comprehensive concept that has been examined in the literature from different perspectives (Lowe et al., 2022; Sodiq et al., 2019). In this study, we have analysed sustainability from the perspective of urban form. Previous studies have analysed sustainability for different urban form/city models, such as compact cities (Bibri et al., 2020), urban sprawl (Egidi et al., 2020), green cities framework (Debrah et al., 2022), transit-oriented development (Knowles & Ferbrache, 2019), and smart cities (Bibri, 2017).

While all these frameworks have some advantages and limitations, there is no consensus as to what constitutes the best urban form from the sustainability perspective (Mobaraki & Vehbi, 2022). Moreover, the applicability of an urban form to a city's planning is influenced by diverse factors including the city's existing land use pattern and resource availability.

Using the sustainability indicators from the different city models and based on our understanding of urbanisation and urban form in Delhi, we identify the following urban form characteristics that can contribute to sustainable urbanisation: high accessibility to services and transit stations, moderate or high built-up density with open spaces, moderate street network density, and large block sizes. Moderate to high built-up density, along with high accessibility to services, ensures spatial equity in the distribution of public services across neighbourhoods. Moderate street network density in areas of high density and high accessibility makes the area compact, which encourages active forms of transportation and reduces street traffic congestion and vehicular emissions (Zhao et al., 2020; Gul et al., 2020; Aram et al., 2019). The study also recognizes the potential benefits of large block sizes in the context of Delhi, where block sizes are generally smaller than the standard norms. Therefore, the inclusion of large block sizes has been deemed necessary for sustainable urbanisation in Delhi.

The first typology (T1) is of high accessibility to services. Along with high accessibility, cluster T1 also has moderate street intersection density and larger block sizes, making it less congested despite its high built-up density. Such characterisation resembles the features of a planned neighbourhood, which is known to provide a better quality of life to its residents (Sharifi et al., 2021; Türkoğlu et al., 2019). In this context, we classify T1 under sustainable urbanisation. From a different perspective, high accessibility to services in planned neighbourhoods also leads to increased housing demand, establishing upscale gated communities and excluding low-income households (Silva et al., 2020; Padeiro et al., 2019).

In Delhi as well, such areas have a very high residential land price per sq. km, approximately four times higher than the average land price (Government of Delhi, 2014). As these areas of high accessibility are primarily populated by high-income households, there is a need for inclusive policies such as (a) affordable housing subsidy for low-income households, (b) inclusive zoning, where a certain percentage of new housing developments are reserved for low- and middle-income households, (c) Building community land trusts, these are non-profit organizations that hold land and make it available for affordable housing or community development purposes (Lowe et al., 2022). Such measures can ensure sustainable urbanisation with social equity.

The second urban form typology (T2) exhibits a moderate level of built-up density and features large block sizes, indicating its potential to facilitate sustainable urbanisation. However, this cluster is also characterised by low street intersection density and limited accessibility, which degrade its sustainability levels. In urban planning, low street intersection density is associated with low walkability and limited access to public transportation, as well as decreased social interaction and community connectivity, which studies have found to have an adverse impact on individual physical and mental health (Baobeid et al., 2020; Xue et al., 2020). Thus, better street design and improved access to transit services are crucial for sustainable urbanisation (Bibri, 2021).

Some policy and planning measures to enhance the street intersection density and walkability can be: (a) Tactical Urbanism: These are low-cost interventions like street paintings, parklets, and pedestrian plazas that can help test new intersections and street designs before committing to permanent changes, (b) Transit-Oriented Development (TOD) to promote the development of mixed-use, high-density developments around public transit stations, which can help increase the density of intersections in those areas, and (c) Narrower Streets: Reducing the width of streets can encourage slower traffic and can also create space for new intersections and crossings (US EPA, 2017).

The third urban form typology (T3) is of newly urbanised areas with a rapid growth rate over the last ten years. Such residential areas also have moderate building density and moderate street intersection density, which aid in sustainable urbanisation. However, low access to services in such areas is a cause for concern. Low accessibility to services diminishes growth opportunities and degrades the quality of life, as previous studies show (Guida et al., 2021, Mouratidis, 2021). To sustain growth and promote sustainable urbanisation in such areas, accessibility to different services needs to be enhanced. One of the important planning interventions in this regard can be encouraging land use diversity in this cluster, which can provide a variety of services and amenities within a single building or block (Pozoukidou & Chatziyiannaki, 2021)

The fourth urban form typology (T4) is of urbanisable areas. These settlements are located on the city's outskirts in isolated pockets and have rural characteristics. Despite their growth over the past decade, they have a low building density. Due to unplanned street networks and block design, features such as large block sizes and low street intersection density are common here. Access to services is also low due to their peripheral location. As these settlements transform from rural to urban, there is an opportunity to improve sustainability through strategic land use and accessibility planning interventions, such as compact and mixed land use development and designing the streets and block size in a manner that promotes walkability and reduces traffic congestion (Alawadi et al., 2022).

The fifth urban form typology (T5) is of compact and congested areas with high street intersection density, small block sizes, and high built-up density. Neighbourhoods with such characteristics, despite having a high degree of interconnected street networks that enhance walkability, also face issues of traffic congestion and environmental pollution (Lu et al., 2021). To achieve sustainable urbanisation in such areas, the following planning interventions can be carried out: (a) promoting the efficient use of land through land conversion policies to create more green spaces; (b) promoting smart mobility through the use of technology to get real-time traffic information; and (c) other measures such as encouraging carpooling and road pricing can reduce the number of cars on the road and alleviate traffic congestion (Dulal, 2017).

The final urban form typology (T6) can be considered similar to that of cluster T5 in terms of high built-up density. However, cluster T6 has a lower street intersection density and higher block size, as compared to cluster T5, which makes it less compact and congested. To achieve sustainable urbanisation, planning interventions of similar nature as highlighted for cluster T5 can be carried out.

After discussing the different typologies, we are now in a position to evaluate the level of sustainable urbanisation in Delhi. The above analysis shows that while a certain urban form characteristic may contribute to sustainable urbanisation in isolation, sustainability within a specific cluster must be evaluated by considering the contribution of each element. In many clusters, there are urban form characteristics that support sustainable urbanisation, such as a large block size in cluster T2 or moderate street intersection density in cluster T3. However, to achieve sustainability in a given cluster, all the urban form elements must have a positive impact. As our results show, only in cluster T1 do all urban form characteristics contribute to sustainable urbanisation.

Considering that the total area of grid cells that come under cluster T1 is only 19% of the total residential area in Delhi (refer to Table 5.3), we conclude that only 19% of the residential area in Delhi can be considered under sustainable urbanisation. The rest of the area requires different forms of intervention to make urbanisation sustainable, as noted in the above paragraphs.

The findings presented in this paper serve as concrete evidence to illustrate the prevalence of urban informality and to provide validation for the theory of spatial mobility in cities in the Global South, such as Delhi. The varied urban features observed in Delhi, including unplanned colonies and slums, serve as tangible manifestations of informality within the urban landscape. Notably, urban theories originating from the Global North have largely overlooked the concept of urban informality as a means of shaping urban spaces—a perspective that many postcolonial urban theorists consider a critique of Western planning practices.

Urban space informality is a complex and multifaceted concept that pertains to various aspects of informal development, occupation, and usage of urban spaces within cities. Several theories and perspectives have emerged to help understand and analyse this phenomenon, for example – (a) Informal Urbanism Theory: This theory focuses on the informal practices, settlements, and activities that emerge in urban areas, often in response to rapid urbanization and inadequate formal planning and infrastructure. Informal urbanism encompasses informal settlements (slums), street vending, unauthorized construction, and other activities that exist outside the formal regulatory framework. (b) Spatial Justice Theory: It argues that informality often arises due to spatial inequalities and injustices, where marginalized communities are excluded from formal urban systems and infrastructure. (c) Supportive Neglect: This concept suggests that governments or authorities may intentionally or unintentionally tolerate informality because it provides a cost-effective solution to urban challenges, such as housing shortages. The term is often used in the context of slums and unplanned settlements.

As our findings show the existence of unplanned neighbourhoods in Delhi can be explained by the theories of urban informality. The ad-hoc nature planning approach results in the creation of "uneven geographies" and contributes to the fragmentation of urban development through the informal shaping of urban spaces, as elucidated by Roy (2009c). Consequently, urban informality is not merely an anomaly but rather a prevailing norm in Global South cities, emphasizing the need for its incorporation into mainstream urban theory and policy discussions.

As highlighted in the Introduction section, this study is important considering the rapid pace of urbanisation in cities in the global south. Given the nature of urbanisation in cities like Delhi, as they become more urbanised, they face various challenges for sustainable development. Noting the impact of urban forms on sustainability, as widely recognised in the literature, this study argues for incorporating the built-up form into the characterisation and measurement of urbanisation. Such an approach, as demonstrated in this study, can provide a more accurate assessment of urbanisation and help bring localised planning interventions to areas that have unsustainable urban form features.

5.6 Conclusion

The study aimed to explore residential built-up form typologies and assess their impact on sustainable urbanisation in Delhi. Only a few studies have explored variations in the urban form at the neighbourhood level, and none exist specifically for cities in the Indian subcontinent. The study used a grid-based technique to divide residential areas into 100 x 100 metre grid cells and assigned attributes of accessibility, built-up density, and street design. The grid cells were then clustered using the k-means clustering algorithm, which showed the presence of six built-up form clusters in Delhi. Using the MANOVA test statistics and graphical visualisations, these clusters were analysed for variation in their urban form elements and were found to be significantly different from one another.

Using the SHAP tool, the clusters were analysed for their dominant urban form, using which cluster typologies were framed. These typologies can be listed as: (1) areas with high accessibility to services; (2) areas with moderate built-up density; (3) areas with a high growth rate; (4) urbanisable areas with a rural landscape; (5) compact and congested areas; and (6) high density areas. The study then discussed how the different built-up form elements in these clusters contribute to sustainable urbanisation in Delhi. Based on the results, the study concludes that only 19% of residential areas in Delhi can be classified under sustainable urbanisation, while the remaining areas require different planning interventions to achieve sustainable urbanisation.

We note here the limitations of the study. First, the study considers only the physical elements of urban form and their association with sustainable urbanisation. The understanding of sustainable urbanisation can be further enriched by including socioeconomic and demographic indicators of neighbourhoods, such as population density, economic status, and age and caste-wise composition. Furthermore, sustainability can be analysed using local environmental and ecological indicators, such as the air quality index. In this manner, one can develop a more comprehensive understanding of sustainable urbanisation in different neighbourhoods, and appropriate localised policies can be developed.

Second, the study has used the cumulative measure of accessibility, which is a potential measure. Other accessibility measures based on time/ distance or demand and supply of services, such as the 2-step floating catchment area method, can be used. This can provide a more realistic measure of accessibility. Finally, with the advancements in computationally efficient learning algorithms, future studies can perform a comparative analysis among different clustering techniques to examine the method with the highest clustering efficiency and utilise it to cluster the urban form at the neighbourhood level.

Despite the shortcomings, the study offers a viewpoint for understanding urbanisation in rapidly urbanising cities like Delhi, which are characterised by spatial heterogeneity in their urban form. Our study shows that urban form at the neighbourhood level can show significant spatial variation, and thus characterising a city with a particular urban form can be misleading. Our study methodology is generalisable to other cities and can be utilised to create development zones that are based on the dominant built-up types and are defined by adaptive and dynamic boundaries. Based on the dominant characteristics of the zones or clusters, the required planning interventions can be sought. Thus, by factoring in the neighbourhood’s built-up form in the analysis of sustainable urbanisation, this study provides another perspective to study urbanisation, on which future studies can build by considering non-physical characteristics of neighbourhoods.

Chapter 6

City Affordability and Residential Location Choice: A Demonstration Using Agent Based Model

Published as: Marwal, A., & Silva, E. A. (2023). City affordability and residential location choice: A demonstration using agent based model. Habitat International, 136, 102816. https://doi.org/10.1016/j.habitatint.2023.102816

Chapter Overview: Who lives where and why? – is a prominent issue studied in the field of urban economics. With urbanization on the rise, housing and transportation policies must strive to strike a balance between accessibility and affordability. The study builds an economic rational agent-based model, for a hypothetical monocentric city, to simulate the urban pattern that emerges from households' residential location choice as they aim to minimize their expenditure on rent and commute under different scenarios. The model highlights the significance of housing and transportation costs as a spatial policy tool in shaping urban growth. By manipulating these costs, cities can promote compactness, increase affordability, and result in a more homogeneous density and income distribution pattern. The study also finds that mode of travel plays a crucial role in determining residential choice, with private transportation users tending to reside in the city's inner areas and public transportation users opting for outer areas. However, when public transportation is heavily subsidized, this pattern is reversed. We also find that an increase in income inequality and plot size variability can lead to income-based segregation in the city. Our study findings, validated through a review of the relevant empirical literature, provide valuable policy directions into the underlying mechanisms that shape the urban growth pattern.

6.1 Introduction

The choice of a household's residential location is influenced by a multitude of factors, including the residential density (Duranton & Puga, 2020), access to public amenities and employment opportunities (Baraklianos, 2020), the neighbourhood's physical and natural characteristics (Mouratidis & Yiannakou, 2022), the cost of housing and plot size (Huang & Chen, 2022), and proximity to similar communities (Guidon et al., 2019). Despite a household's desire to optimize all of these factors, budget constraints often limit their options (Stephen Ezennia & Hoskara, 2019). According to the 2020 Consumer Expenditure Survey, housing and transportation expenses account for the largest portions of a household's budget in the United States, representing 15% and 12% of their income, respectively (U.S. Bureau of Labor Statistics, 2021). While these two expenditures are significant, they are also interrelated, with higher accessibility to employment opportunities often corresponding to higher housing costs and lower commuting expenses (Haas et al., 2016). As households are bound by a fixed budget, they must make trade-offs between housing and transportation costs in order to maximize their overall utility (Rehman & Jamil, 2021).

As each household aims to maximize its utility, the interplay between demand and availability of preferred residential locations leads to the formation of diverse urban patterns, including slums and informal settlements (Basso et al., 2021). Studies show that land use regulation policies also impact the formal-informal divide in residential areas (Heikkila & Harten, 2023). The evolved urban patterns can be characterized by several parameters such as residential density, land-use mix, plot size, residential segregation, and income distribution. Research has found that these parameters vary with distance from the city centre and between affluent and disadvantaged neighbourhoods (Zhao et al., 2018; Gelormino et al., 2015). Moreover, the variation exists across cities of different sizes, populations, and land-use policies. While the impact of housing and travel affordability on household location choices and the resulting urban morphology has been analysed through various economic models, only a few studies have simulated these relationships (Yen et al., 2019).

The study builds an economic rational agent based model to simulate the urban pattern that emerges from households' residential location choices as they make decisions to minimize their expenditure on rent and commute under different scenarios. The scenarios are created by varying housing and transportation cost, income inequality, mode of travel, and occupied area. The study sets three objectives: (a) to examine the trade-off between rent and commuting expenditure as reported in traditional location choice models, (b) to understand how to make residential density more homogeneous across the city to avoid overcrowding and sprawl, and (c) to investigate the impact of unequal income and land ownership distribution on residential segregation. It is important to note that the model only considers rent and commuting expenses in determining household location choices. While other factors such as proximity to friends and family, neighbourhood vibrancy, and access to green areas may also impact location choice, this study assumes households to be rational economic agents who aim to maximize their utility by minimizing expenditure. Adding more factors to the model would complicate the model design and operation and detract from the study's aim of simulating the possible urban pattern that may emerge as cities become more or less expensive for households.

The study's major contribution lies in simulating some of the complex phenomena of urban development concerning city size and residential segregation. With an understanding of how the residential location choice of different income households varies with city affordability, appropriate policies can be made to solve these issues and promote more equitable urban development. By incorporating various factors such as housing and transportation costs, income distribution and land ownership in a city, the simulation results provide valuable insights into the underlying mechanisms that shape the urban growth pattern. This can inform policymakers and urban planners to make data-driven decisions aimed at improving the livability and sustainability of cities.

The structure of the paper is as follows: Section 6.2 provides a brief literature review on household expenditure and location choice. In Section 6.3, we outline the design and calibration of the Agent-Based Model (ABM). In Section 6.4, we present the results from the ABM simulation. Section 6.5 validates and discusses the findings of the model, and finally, Section 6.6 concludes the paper.

6.2 Literature Review

The Sustainable Development Goal (SDG) 11 aims to make cities more resilient and sustainable by providing better access to public transportation and affordable housing for all. However, current practices in urban policy and planning tend to address these issues separately instead of in an integrated manner (Nakamura & Avner, 2021; Waddell et al., 2003). Research shows that improving access to public amenities in a neighbourhood can lead to higher rent prices, making housing unaffordable for low income households (Brueckner et al., 1999; Corrigan et al., 2019; Hamilton & Röell, 1982; So et al., 2001). The trade-off between rent price and commuting cost has been explored in examining household residential location choices and average commuting distance. One of the seminal theories in urban economics that conceptualise and model household location choice is the Alonso-Muth-Mills Model (AMM), formulated by Wheaton (1974).

To elucidate the fundamental principles of urban economics governing the intrinsic interplay between land rents and location behaviour, we employ Alonso's bid rent theory (1962). This theory formulates the decision-making process of households regarding their residential location as a delicate balance between proximity to the city centre and the available residential space within a monocentric city featuring an immutable central business district (CBD). In this context, the sole spatial attribute of each city location that holds significance for households is its distance from the CBD. As is commonly assumed in consumer behaviour theory, households are posited to aim for utility maximization while adhering to budgetary constraints. Mathematically, the foundational model for the selection of residential location can be expressed as follows:

$$\max U(s,z), \quad subject\ to\ \ z + R(r)s = Y - T(r) \qquad (6.1(a))$$

$$r \geq 0, s > 0, z > 0$$

Where z represents the amount of composite consumer goods which includes all goods except land, s is the lot size of the house. Y represents the household income that is spent on composite goods, land and transportation. $R(r)$ is rent per unit of lot size at distance r, and $T(r)$ is the total transportation cost at r. The bid rent $\varphi(r,u)$ is the maximum rent per unit of land that the household can pay for residing at a distance r while enjoying a fixed utility level u.

Using Equation 6.1(a), the bid rent can be mathematically expressed as,

$$\varphi(r,u) = \max\left\{\frac{Y - T(r) - z}{s}\,\middle|\, U(z,s) = u\right\} \qquad (6.1\ (b))$$

Graphically, as depicted in Figure 6.1(a), bid rent $\varphi(r,u)$ is given by the slope of the budget line at distance r that is just tangent to the indifference curve u.

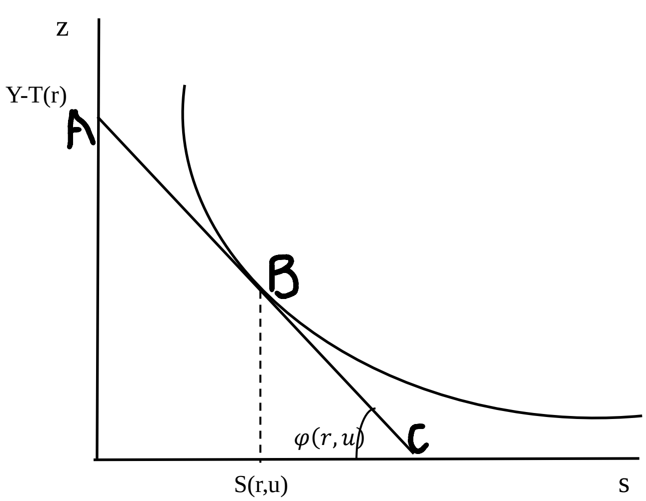

Figure 6.1(a) – Bid rent function

We can conclude that bid rent φ(r,u)), that is, the highest land rent at r under which the household can achieve utility level u, is given by the slope of budget line AC. The tangency point B determines bid-max lot size $S(r,u)$.

Using the bid-rent function, we can determine the equilibrium location of the household under a given land rent configuration of the city. The equilibrium location of the household is that location at which a bid rent curve is tangent to the market rent curve from below. The above analysis from the economic model shows that household residential location choice can be derived by studying the land rent function in a city. The above model can be extended to understand how the location choice changes with variations in commuting cost, wage rate and household family size.

According to the AMM, in a monocentric city, density and rent price decreases with distance from the city centre while household plot area and commuting cost increases. The combined expenses on housing and transportation reach an equilibrium where a decrease in housing price is offset by an equal increase in transportation cost and vice versa. Many other traditional urban economic models also assume this perfect tradeoff between housing and transport expenditure (Mattingly & Morrissey, 2014). However, the empirical evidence from different cities clearly shows that living cost in cities is not the same for all households. It is un-equally affordable for households across space and socio-economic groups (Dewita et al., 2020; Makarewicz et al., 2020). The reason behind this inequality in living costs is largely twofold – income inequality and location efficiency. High income inequality results in poor income households spending a higher percentage of their income on housing and transportation, even at locations where the average expenditure is less (Campbell, 2021; Charles & Lundy, 2013).

Location efficiency is about more sustainable use of resources to cut down wasteful commuting and provide a vibrant environment for living. One crucial aspect of location efficiency is access to public transportation. Poor access in fringe areas often leads to a reliance on private vehicles for daily commuting, adding a significant capital expenditure to the household budget and increasing daily transportation costs (Banister, 1994; Currie & Senbergs, 2007). This not only raises the daily transportation cost but also increases the household's total expenditure-to-income ratio (Viggers & Howden-Chapman, 2011). As a result, households residing in fringe areas, despite the lower housing rent, often incur higher overall living costs than what traditional economic theory suggests (Kellett et al., 2012). However, it is important to note that household location choices may not always be based solely on optimizing costs, as studies have shown that personal attitudes, beliefs, and lifestyle preferences can also play a role in their decision making, leading to higher expenditures on both housing and transportation (Deka, 2015).

From the above discussion, it is clear that owning to location efficiency, household economic status, and their subjective considerations, households make unequal expenditures on housing rent and commuting. This implies that households tend to compete for locations where they can meet their desired expenditure-to-income ratio. Based on the household residential location decisions the city has variations in population density, commuting distance, land price, house rent, and plot size. These aspects of the city ultimately define the neighbourhood morphology and liveability of the households.

Using ABM as a simulation tool

Simulation can help to comprehend the emergence of city morphology as inhabitants engage with the built environment (Batty, 2007). Currently, agent-based models (ABM) are frequently used to simulate the behaviour of discrete decision-makers in an interactive setting. The simulations, performed by applying simple interactive rules between agents, can reveal the complexity of the system that arises from the collective behaviour of the agents. Furthermore, ABMs, owing to their dynamic nature, can provide insight into complex phenomena that are challenging to analyze using traditional statistical equilibrium models (Yen et al., 2019).

ABMs have been widely applied in various areas of urban planning, such as real estate modelling (Zhuge et al., 2016), transportation (Babakan & Taleai, 2015), economic development (Leao et al., 2017), segregation (Crooks, 2010), informal housing (Patel et al., 2012), disaster risk reduction (Crooks & Wise, 2013) and land use and transportation (Acheampong & Asabere, 2021). However, only a few studies have employed ABMs to simulate the change in residential location choices with variations in travel and housing costs. For example, Kulish et al. (2012) used ABMs to analyze how city morphology changes with changes in land use and transportation policies. Raux et al. (2014) used an ABM model to show the trade-off between commuting costs and rent costs and found results similar to those of traditional statistical equilibrium models. Yen et al. (2019) built an ABM inspired by traditional statistical models to simulate the density patterns under different income distributions, employment centres, transportation modes, and infrastructures.

In the above studies, the decision-making process for household relocation is centred on the maximization of utility through the minimization of rent and commuting costs. However, this approach fails to consider the expenditure levels of varying income households which they are willing to incur. Such a consideration is imperative as it determines the affordability of a location for a household. Location affordability serves as an income-based relative measure that informs the financial viability of a residential location to a household. The issue of affordability assumes greater significance since households tend to choose a location that not only provides them with increased access to resources and lower housing and travel costs but also meets their desired affordability threshold based on their income levels.

In light of this potential gap in the existing literature, this study aims to address the impact of varying levels of affordability on the urban form by constructing an Agent-Based Model (ABM). Specifically, the model will allow for the identification of the resulting urban form as the city becomes more or less affordable for different income groups. By incorporating consideration of affordability into the decision-making process of agents, this study seeks to provide a more nuanced understanding of the interplay between affordability, residential location choices and urban form.

6.3 Model and Parameters

6.3.1 Overview and design

The ABM model is built for a monocentric city in the NetLogo software (Wilensky, 1999). Households, represented as agents, are randomly placed throughout the city landscape and assigned income based on random-normal income distribution. The agents use their income only for housing rent and commuting to the city centre. They are in one of two modes: searching or satisfied. At the start of the model, all agents are in search mode as they search for a location that minimizes their expenses. When an agent finds a suitable location, they become satisfied. The model ends when all agents are satisfied. A conceptual overview is shown in Figure 6.1.

The city is designed as a circular ring of radius of 14 units and a patch area of 10 units. The central patch is non-residential where all amenities and services are located, including agents' place of work. Other than the centre, agents can occupy any patch. However, a patch can house a maximum of 10 agents, with each agent occupying 1 unit area. If a patch reaches its maximum capacity, it is no longer available for housing. Figure 6.2 shows the city landscape. The number of agents (households) generated is set to half of the city's total housing capacity, as calculated by Equation (6.1).

Every patch is characterized by two factors: patch density and proximity to the city centre. Patch density is calculated by adding the equally weighted ratios of satisfied agents in the patch and its Moore neighbourhood (8 neighbouring cells) to the total number of agents that the patch and its Moore neighbourhood can hold. The maximum number of agents in a Moore neighbourhood is 80 (10 agents per patch). Agents calculate their rent expenditure using Equations (6.2) and (6.3). Rent is assumed to be directly proportional to patch demand, expressed in terms of patch density.

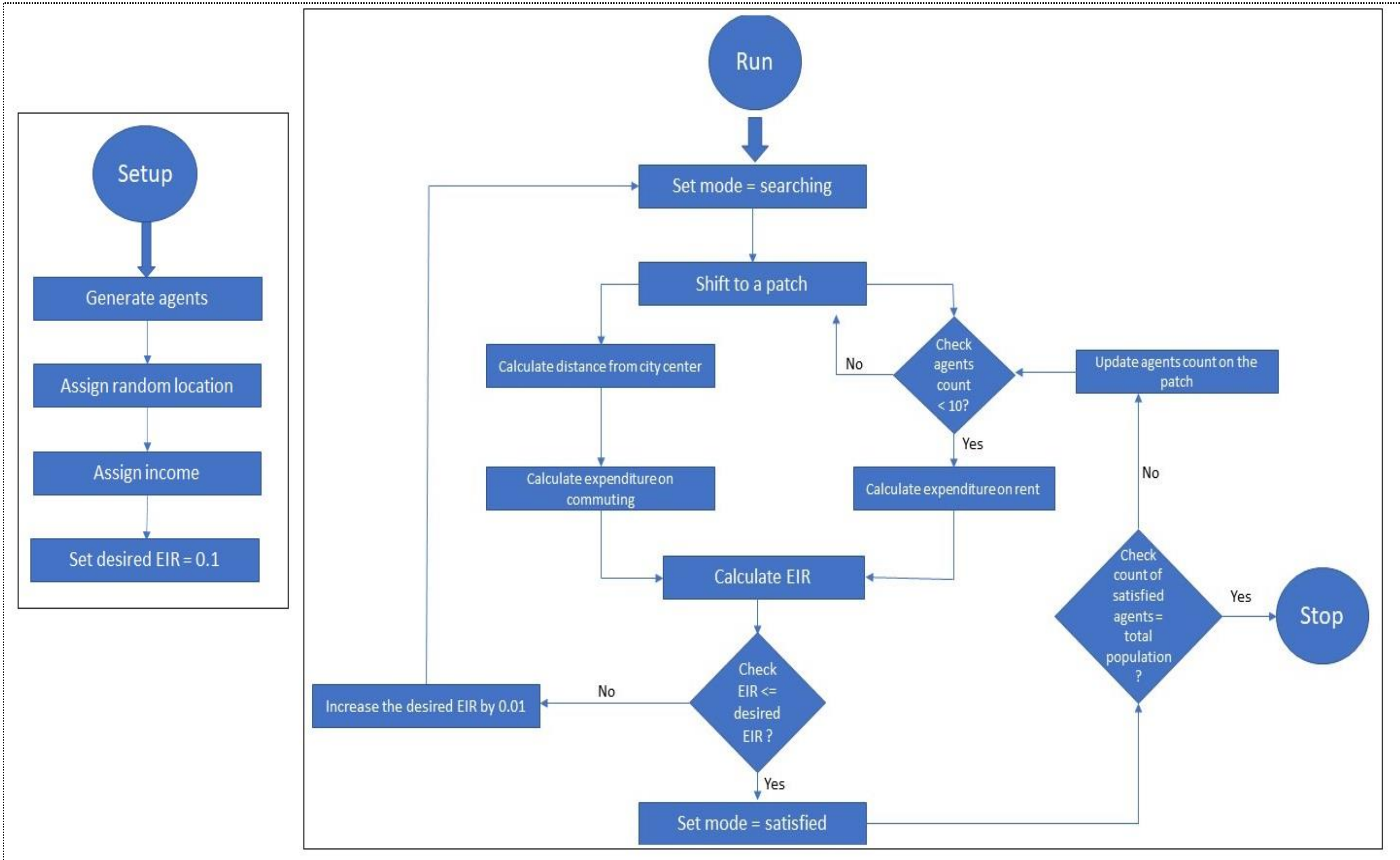

Figure 6.1: Model flow chart

Commuting expenditure (Equation (6.4)) is expressed as a linear function of patch distance from the city centre and assumes the same travel mode and fuel prices for all agents.

At the start of the model, agents are randomly placed on a patch and are in search mode. When they move to a patch, they calculate their expenditure-to-income ratio (EIR) using the patch density and distance from the city centre, as shown in Equation (6.5). Each agent on a particular patch incurs the same expense on rent and commuting, but their EIR may vary due to differences in income.

$$\text{Total agents (N)} = \frac{(\text{Total patches} * 10)}{2} \tag{6.1}$$

$$\text{Patch Density (d)} = \frac{\text{Count of satisfied agents on the patch}}{2*10} + \frac{(\text{Count of satisfied agents on the patch neighbourhood})}{2*80} \tag{6.2}$$

$$\text{Expenditure on rent } (E_r) = \text{r} * \text{d} \tag{6.3}$$

$$\text{Expenditure on commuting } (\text{E}_\text{t}) = \text{t} * \text{Patch distance from the city centre} \tag{6.4}$$

$$\text{Expenditure to Income Ratio (EIR)} = \frac{(E_r + E_t)}{\text{Income}} \tag{6.5}$$

The housing rent coefficient is labelled as "r" and the commuting coefficient is labelled as "t". These coefficients can also be considered as the housing rent price per unit area and the commuting cost per unit distance, respectively.

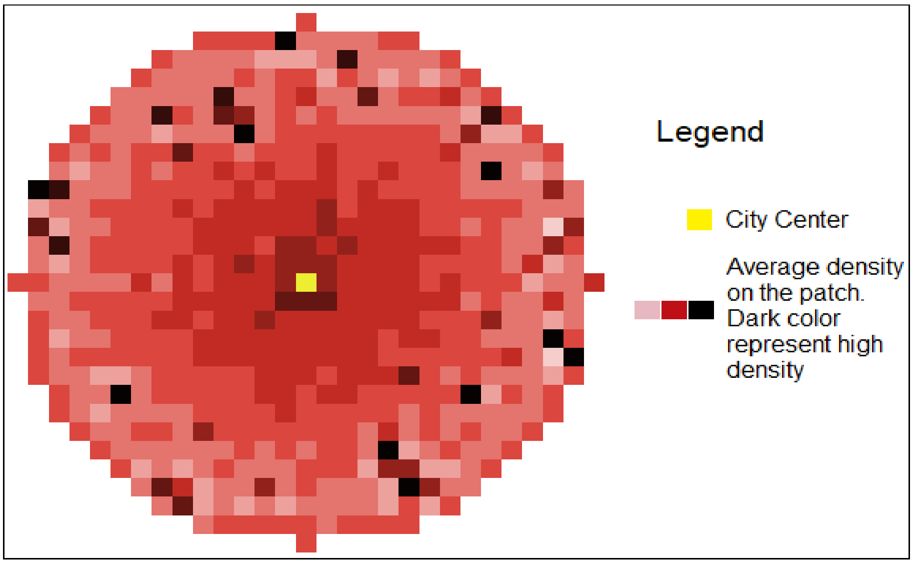

Figure 6.2: City landscape with variation in patch density

6.3.2 Detail and model calibration

The agents in this model are rational economic consumers who aim to minimize their expenditure-to-income ratio (EIR). The minimum EIR an agent can have is referred to as the desired EIR. At the start of the model, all agents have a desired EIR of 0.1 and are in search mode. During each iteration, 10% of the agents in search mode are randomly selected to evaluate their current EIR. If an agent's EIR is found to be equal to or lower than their desired EIR, they are labelled as satisfied. However, if the EIR is higher than their desired EIR, they search for a new location with an EIR equal to or lower than their desired EIR. Each agent can search for up to 10 locations in each iteration. If they fail to find a suitable location, their desired EIR is increased by 0.01 and they continue the search in future iterations. The model allows the unsatisfied agents to keep on searching in an unlimited manner, without having a relocation cost. The search process continues until all agents are satisfied.

The density of a patch and its surrounding neighbourhood changes as agents settle into or vacate the patch. This causes the housing rent on a patch to be proportional to its demand, making the expenditure on rent dynamic. However, the expenditure on commuting remains constant for a given patch. Only satisfied agents are counted in the calculation of patch density and only permanent residents are considered. Agents in search mode are temporarily located on a patch and are not counted in the density calculation.

We first run the standard model, where the agents' income is generated using a random normal income distribution, with a mean income of 50,000 units and a standard deviation of 12,000. The minimum income is set at 10,000 units. Agents are divided into two groups based on their income: those with income higher than the mean income are classified as high income agents, and the rest are classified as low income agents. The input values for the standard model are outlined in Table 6.1. The results of the model help us understand how agent density and EIR vary with the city centre.

After running the standard model, four experiments are performed to evaluate the impact of different factors on the model. The first experiment involves changing the coefficients of rent and commuting to make them more or less expensive than the standard model. The second experiment involves varying income inequality levels using the beta distribution function of income and measuring the Gini coefficient. The third experiment involves randomly assigning two travel modes (public or private) to agents, with different commuting coefficients. The fourth experiment allows agents to occupy more than one unit of area on a patch. The results of these experiments are presented in section 6.4.

Table 6.1: Input variables

Input variables	Mean
Mean income	50000
Mean distance from city cent**re**	7.5
Total agent population	3065
Rent coefficient (r)	30000
Commuting coefficient (t)	700
Minimum expenditure on rent	2000
Minimum expenditure on commuting	700

Model Calibration

To calibrate the model, we begin by randomly assigning income to each agent using a normal distribution with a mean income of 50,000 units per month. In a real-world scenario, income distribution curves are often positively skewed with a high proportion of households having income lower than the mean income. The model assumes income to be normally distributed in the standard model with a low level of income inequality. The selection of the mean income has no impact on the outcome of the model as the model inputs the income-to-expenditure ratio rather than the absolute income. The value chosen for mean income is therefore arbitrary.

To set the values of rent and commuting coefficients, we consider the average household expenditures on these items from real-world scenarios. According to a recent OECD report on household expenditures (OECD, 2023), there is a significant variation in mean expenditures on rent and commuting among countries. For instance, countries like Finland, the UK, Japan, and France have households spending over 25% of their income on housing-related expenses and rent, while countries like Malta, Lithuania, and Turkey report expenditures below 15%. Commuting expenses tend to be more consistent across countries, with an average of around 10%. Based on these findings, we set the average expenditure on household rent and commuting to 20% and 10% of income, respectively. With a mean income of 50,000 units per month, this results in a mean expenditure of 10,000 units per month on rent and 5000 units per month on commuting. The coefficient of rent (r) can be calculated using Equation (6.3), as shown in Equation (6.6):

$$r = \frac{\text{Expenditure on rent } (E_r)_{mean}}{d_{mean}} \tag{6.6}$$

Where d_{mean} denotes the fraction of residential area in the city. The ideal percentage of the land area allocated for residential purposes in a city can vary greatly depending on a number of factors, such as population density, land availability, infrastructure and transportation networks, local zoning laws and regulations, and cultural and economic factors. There is no universally agreed upon ideal percentage, as the needs and priorities of different cities can differ greatly. The value of the mean residential density d_{mean} in our model is based on the assumption that the ideal land use classification in a city should allocate 1/3rd of the area to green spaces, 1/3rd to industrial, infrastructure, and commercial uses, and the remaining 1/3rd to residential purposes. Using Equation (6.6), the rent coefficient (r) is calculated as 30000 per unit area.

Solving Equation (6.4), we get the expression for commuting coefficient (t) as shown in Equation (6.7):

$$t = \frac{\text{Expenditure on commuting } (\mathrm{E_t})_{mean}}{\text{Patch distance from the city centre}_{mean}} \tag{6.7}$$

As determined above, the mean expenditure on commuting $(E_t)_{mean}$ is 5000 units and the mean patch distance from the city centre in our model is 7 units, which makes the value of the commuting coefficient (t) nearly equal to 700 per unit distance.

6.4 Results and Analysis

6.4.1 Standard model

This model tests the empirical relationship between variation in urban density with distance from the city centre in a monocentric form. The key output parameters are listed in Table 6.2. As shown in Figure 6.3(a), the agents' density decreases linearly as the distance from the city centre increases. The coefficient of variation (CV) for density is 0.17, indicating a low variation in density with distance from the city centre. This leads to a low variation in rent as well, meaning that rent is less sensitive to the distance from the city centre. To minimize total expenditure, agents are more likely to reside close to the city centre to reduce commuting costs. The variation in income with distance from the city centre reveals that the high income group tends to occupy the intermediate areas between the inner and outer regions of the city, while agents with lower income settle within the inner and outer fringes of the city (Figure 6.3(b)).

Table 6.2: Key output variables in the standard model

Output Variables	Mean	Minimum	Maximum	CV
Density	0.5	0.31	0.7	0.17
EIR	0.3	0.1	0.8	0.63
E_r	6978.69	2000	19125	0.6
E_t	6241.81	700	9800	0.38
Total Expenditure	13220.5	2700	21018.58	0.32

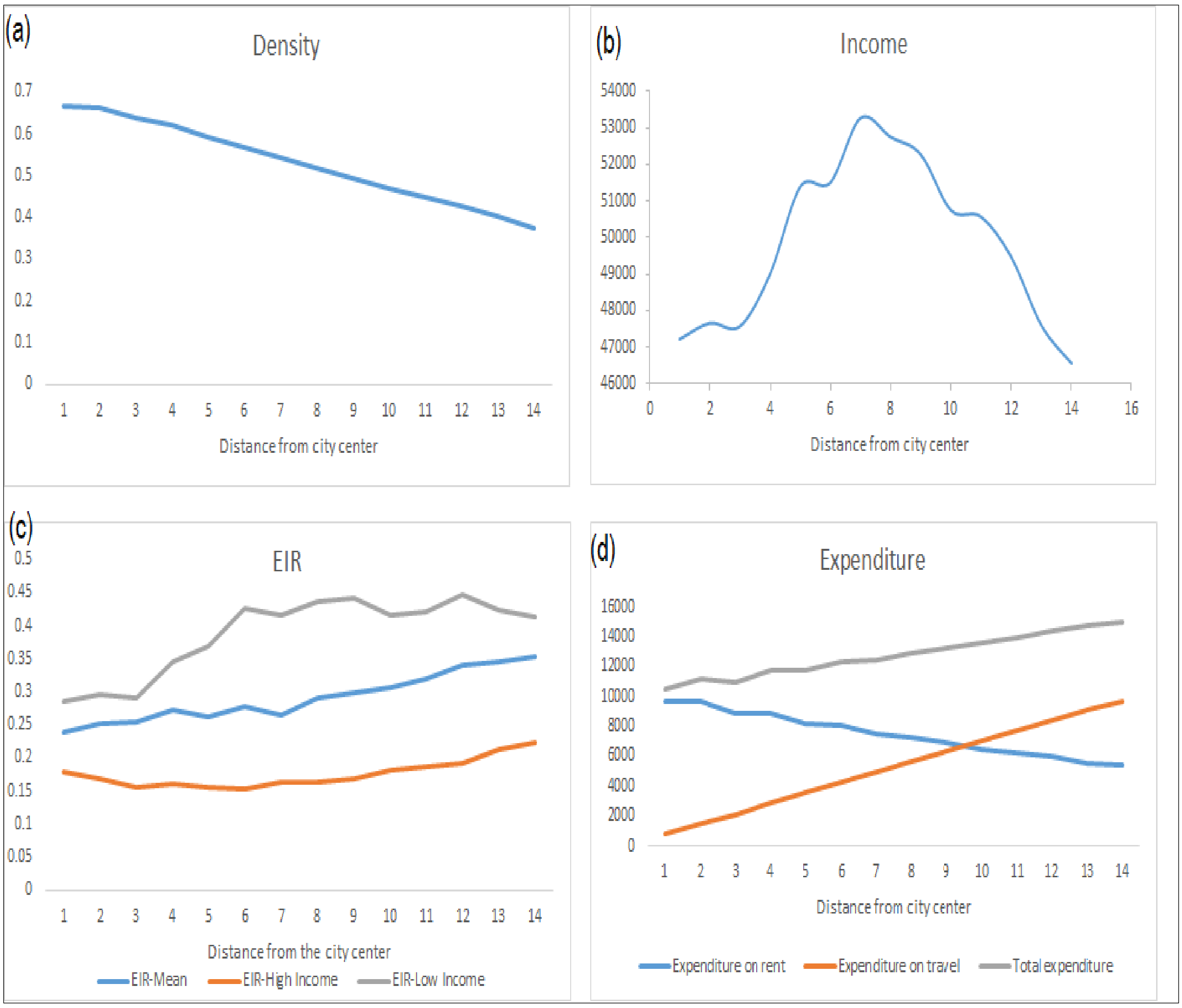

Figure 6.3: Variation in (a) density, (b) income, (c) EIR, and (d) expenditure with distance from the city centre.

We now look at the variation in EIR with distance from the city centre. Figure 6.3(c) demonstrates a linear increase in the EIR as the distance from the city centre increases, with a range of 0.23–0.35. This indicates that living costs, including rent and commuting expenses, increase as one moves away from the city centre. It's noteworthy that the EIR differs between high and low income groups. On average, high income agents spend 40% of their income on rent and commuting, while low income agents spend 17%, resulting in a 23% higher expenditure on these expenses for low income agents as compared to high income ones, according to our model. We also notice how expenditure on rent declines and expenditure on commuting increases with an increase in distance from the city centre (Figure 6.3(d)). However, the decline in one is not completely offset by the increment in another, which makes the total expenditure increase with an increase in distance from the city centre.

6.4.2 Experiment 1: Variation in the rent and commuting coefficient

In this part, we investigate how changing the rent and commuting coefficients affects location choice and city density. We use 9 combinations of rent and commuting coefficients, obtained by choosing coefficients 50% higher and lower than their standard values, and run the model to observe the variation in density pattern and location choices (as shown in Table 6.3). The density patterns are illustrated in Figure 6.4(a). The lowest variation in density occurs when the rent coefficient is higher (r_3) and the commuting coefficient is lower (t_1) than their mean values. Conversely, the highest variation in density occurs when the rent coefficient is lower (r_1) and the commuting coefficient is higher (t_3) than their mean values.

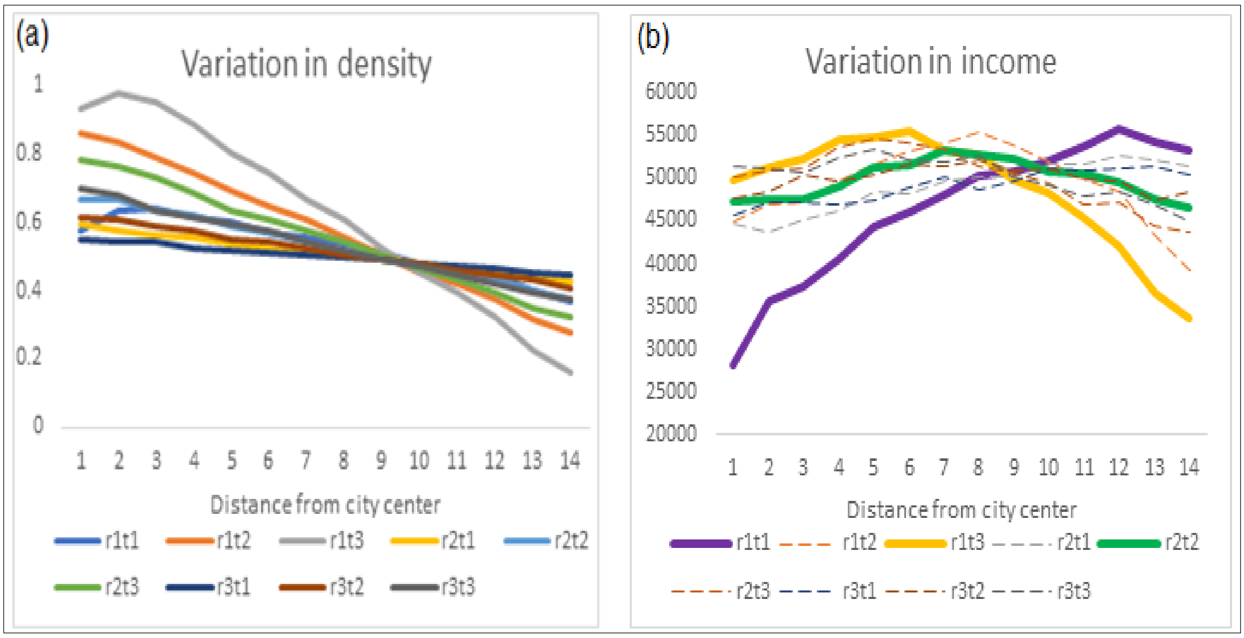

Figure 6.4: Variation in (a) density, and (b) income with change in rent and commuting coefficients.

Table 6.3: Combinations of rent and commuting coefficients

r/t	350	700	1100
15000	r_1t_1	r_1t_2	r_1t_3
30000	r_2t_1	r_2t_2	r_2t_3
45000	r_3t_1	r_3t_2	r_3t_3

When the commuting coefficient is low, a greater distance from the city centre has a smaller effect on commuting costs. If the commuting coefficient remains low and the rent coefficient is increased (combination r_3t_1), agents strive to minimize their EIR by living in less dense areas, leading to urban sprawl. Conversely, if the rent coefficient is low, the cost of rent is less sensitive to changes in density. In this scenario, with a high commuting coefficient (combination r_1t_3), agents opt to live close to the city centre, causing overcrowding in inner city areas.

The spatial distribution of high and low income agents is also worth noticing. For the scenario with the lowest rent coefficient (r_1), two distinct income distribution patterns emerge. When the commuting coefficient is high (t_3), low income agents live away from the city centre. Conversely, when the commuting coefficient is low (t_1), high income agents reside away from the city centre, creating patterns of economic segregation. This pattern repeats as the rent coefficient increases, although with less variation in income levels as a function of distance from the city centre, as seen in Figure 6.4(b). This suggests that when commuting is more affordable, high income households prefer to live in outer city areas, spending less on rent, while low income households are pushed towards the city centre.

6.4.3 Experiment 2: Variation in income inequality

In this experiment, we analyse the change in location choice and EIR for high and low income groups with a change in income inequality. Using beta distribution, we generate two income distributions having a Gini coefficient of 0.3 and 0.45. With increasing income inequality, the majority of agents' income decreases, leading to an increase in the mean EIR

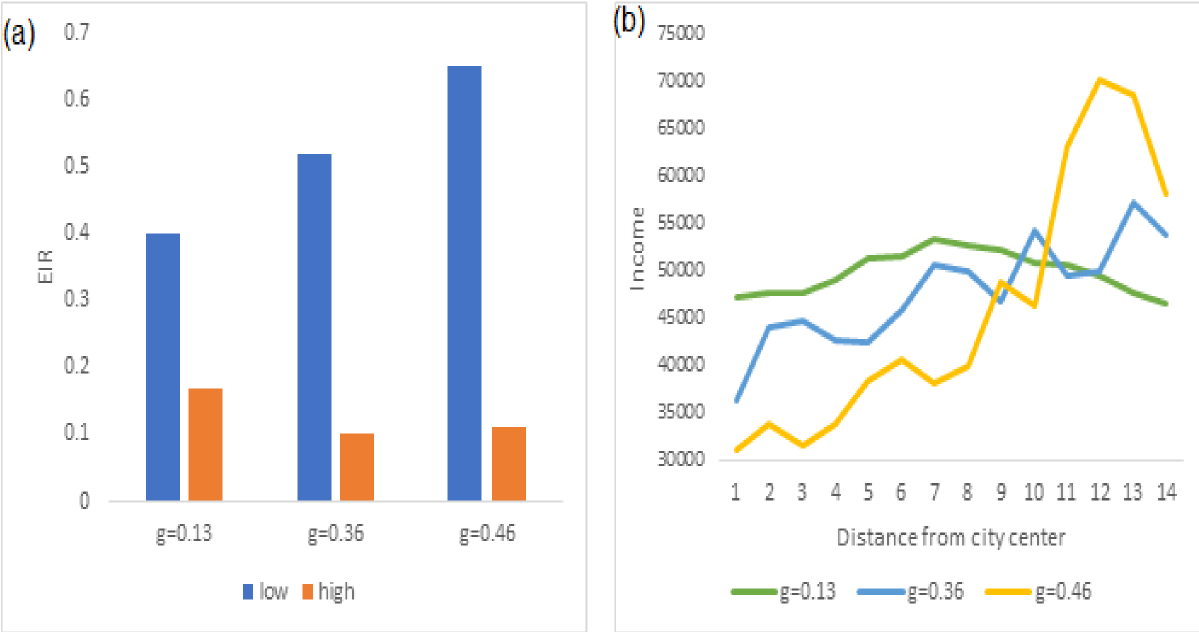

Figure 6.5: (a) Variation in mean EIR for low and high income groups with income inequality. (b) Variation in income with distance from the city centre for different income inequality.

The average EIR rises from 0.3 to 0.6 as income inequality increases from 0.13 to 0.46. It is noteworthy that with an increase in income inequality, the EIR for the high income group decreases while that of the low income group increases, as shown in Figure 6.5(a). This indicates that high income agents spend a smaller proportion of their income on rent and commuting when income inequality is high. Another interesting pattern observed is that as income inequality increases, the high income group is more likely to be located away from the city centre, while the low income group comes closer to the city centre, as shown in Figure 6.5(b).

6.4.4 Experiment 3: Variation in travel mode choice

In this experiment, we examine the impact of travel mode on residential location choice. Agents are randomly assigned to either public or private transportation modes. The commuting coefficient for the private mode of travel is fixed at t_1=1000, while the commuting coefficient for the public mode of travel, t_2, is varied from 700 to 100.

At t_2 equal to 700, 500, and 300, a higher proportion of agents using public transportation reside outside the city centre, more than 10 units away (see Figure 6.6 (a)). However, as t_2 decreases to 100, the trend changes and a greater number of agents using private transportation now locate in the outer city area, as shown in Figure 6.6 (b). This is a very interesting finding as it shows that making public transportation less and less expensive can significantly alter the residential location choice of commuters, drawing public transportation users closer to the city centre

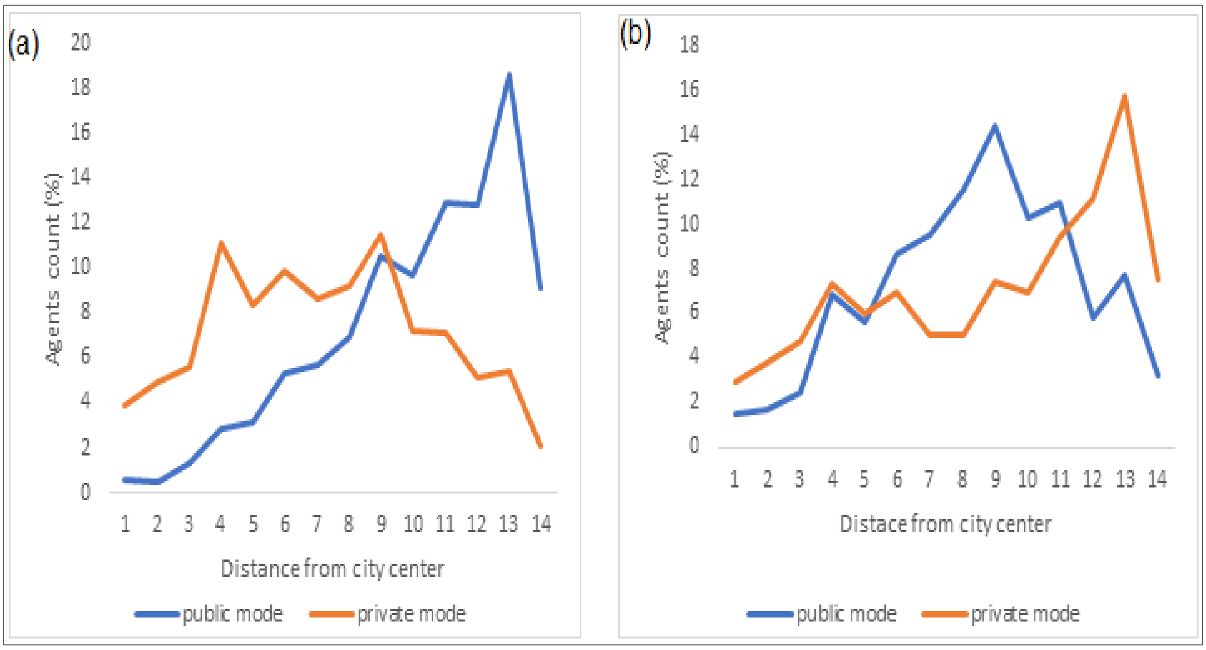

Figure 6.6: Variation in agents' count (%) with distance from the city centre using public and private modes of travel for (a) t_1=1000, t_2 = 500, and (b) t_1=1000, t_2 = 100

6.4.5 Experiment 4: Variation in household occupied area

In our final experiment, we investigate the variation in household plot area with distance from the city centre. Similar to the standard model, in this experiment each patch has a total area of 10 units. However, now agents can occupy between one to two units of area per patch. As an agent moves to a patch, it determines the desired area to occupy based on its budget and the patch's per unit rent, which is calculated as the rent coefficient multiplied by the patch density (Equation (6.8)). If the patch has enough available space, the agent settles, otherwise it moves to another patch. After 10 attempts without finding a suitable patch, the agent increases its desired EIR by 0.01. To make the model more realistic, we set a minimum desired area of 1.5 units for high income agents. The results show a linear increase in the mean area occupied by agents with increasing distance from the city centre.

$$Desired\ area = \frac{(Desired\ EIR * Income - Expenditure\ on\ commuting)}{Rent\ per\ unit\ area} \tag{6.8}$$

Studies have found that in monocentric cities, the average household plot size increases with distance from the city centre, due to lower density and rent. Our model results also show that when high income agents prefer larger plots, agents' income and the occupied area on a patch increase with an increase in distance from the city centre (as shown in Figure 6.7). The change in income distribution from a bell-shaped curve in the standard model (Figure 6.3(b)) to a linear curve indicates that high income agents shift from the intermediate zone to the outer area to occupy plots of a bigger size. The mean plot size occupied by agents was found to be 1.38 units, exhibiting a 25% increase from the inner to the outer city. We also studied the impact of varying the rent and commuting coefficients on the occupied area however, the changes were found to be minimal.

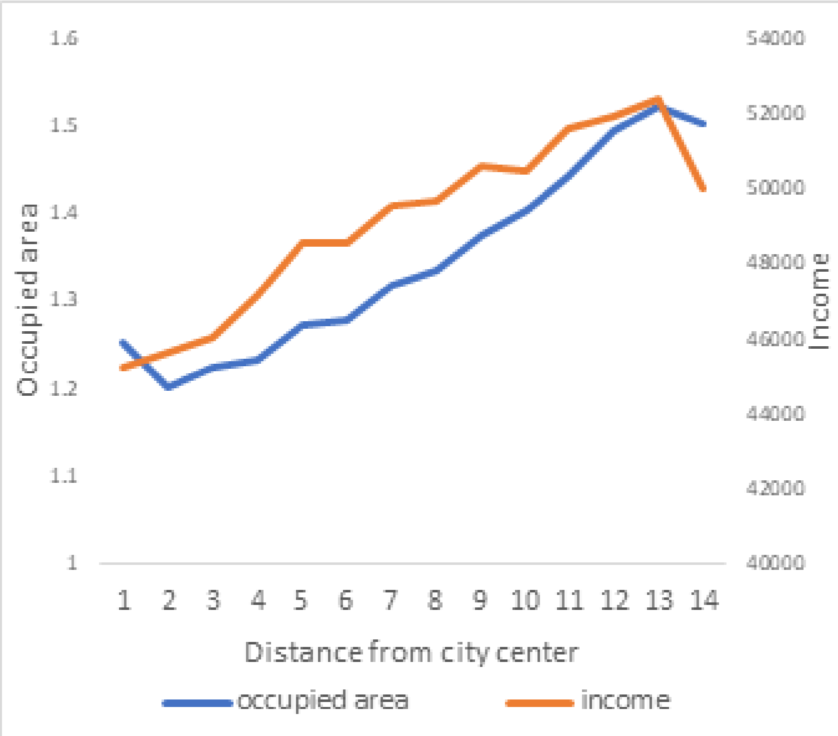

Figure 6.7: Variation in the occupied area and income from distance from the city centre.

6.5 Result Verification and Discussion

The paper performed an ABM simulation to study the impact of city affordability, measured in terms of expenditure on rent and commuting, on agents' residential location choices. Through various experiments, the location choices of high and low-income agents and the resulting density pattern were analysed. The validity of the model's main results will now be assessed through a review of the relevant empirical literature.

The results of our standard model indicate a linear decline in city density with increasing distance from the city centre, which is a pattern commonly observed in many cities across the globe. For example, a spatiotemporal analysis of urban land densities by Xu et al. (2019) found this pattern in multiple cities in the USA, such as Minneapolis, Los Angeles, and Houston. Similar results have been reported for cities such as Manchester and Berlin (Dong et al., 2019).

The second observation from our model is a bell-shaped income curve as a function of distance from the city centre, suggesting that high income households are more likely to reside in the city intermediate zone. This trend has been noted in several French cities, such as Paris and Lyon, and

North American cities like New York and Chicago (Lemoy et al., 2013). Studies by François et al. (2011), Caubel (2005), and Glaeser et al. (2008) have also confirmed this pattern, where the city's intermediate area is occupied by wealthy households while the inner and outer areas of the city are populated by lower income households.

The third observation pertains to the variation in the expenditure-to-income ratio (EIR) with distance from the city centre. Our model shows an EIR between 25% and 35%, which is a common benchmark for households (Gabriel et al., 2005), and an increase in EIR with increasing distance from the city centre. This suggests that housing affordability is higher in the city's inner area compared to the outer area. Studies by Saberi et al. (2017) in Melbourne, Mattingly and Morrissey (2014) in Auckland and Kellett et al. (2012) in Adelaide have all reported similar results, where the suburbs are found to be less affordable than the inner city when housing and travel costs are considered. Our model also shows that the increase in travel costs with distance from the city centre is not offset by the decrease in housing rent, resulting in households in the outer areas having a higher expenditures to income ratio than those in the inner areas. This trend has also been noted by Liu et al. (2021) in their study of housing and transport costs in the Chicago metropolitan area, where they found that households in the inner city have lower housing and transportation costs compared to those outside the central city.

The results of experiment 1 in the paper provide valuable insights into the impact of transportation policies on urban development. The experiment shows that as the commuting cost increases, the city becomes more compact and less sprawl. This highlights the significance of transportation policies in shaping urban spatial planning. De Vos and Witlox (2013) conducted a study in the Flanders region of Belgium and found that cheap and convenient transportation has led to overconsumption of travel and urban sprawl, which can be addressed by increasing the cost of transportation, particularly for car users. Similarly, Lennox (2020) found that in cities in Australia, due to work-from-home policies, the frequency of travel has decreased, resulting in a flatter density gradient and urban sprawl. These findings further validate the results of experiment 1 and highlight the importance of considering transportation policies in urban development.

Another important observation from the results of experiment 1 is the increase in sprawl with the rise in rent prices. While many studies have explored the variation of property or housing prices with distance from the city centre, the impact of housing prices on location choices has received less attention. Only a few studies have examined this relationship empirically. For example, So et al. (2001) in their study of Iowa, USA found that a rise in housing costs decreases metropolitan residency and increases non-metropolitan residency. Ahrens and Lyons (2021) found that in the Dublin metropolitan area, an increase in rent is associated with an increase in commuting time, indicating increased demand for housing in suburban areas.

The results of experiment 2 indicate that with an increase in income inequality, high income agents tend to move from the middle part of the city to its outer area, while low income agents move towards the city's inner and middle areas. This shows that income inequality exacerbates the segregation between high and low income groups. This is a crucial finding for inclusive urban development. The findings are supported by studies in the literature that show that income inequality is a significant factor in socio-economic segregation (Musterd et al., 2017). Tammaru et al. (2020) found a positive relationship between income inequality and residential segregation in European cities from the 1980s–2000s. Quillian and Lagrange (2016) also found that income segregation in US cities was driven by increasing income inequality, with low income households concentrated in the city’s inner area and high income households in the city suburbs, consistent with the results of our model.

The result of experiment 3 reveals the relationship between location choices and mode of transportation. As previously mentioned, increasing the cost of transportation can lead to a more compact city. This experiment supports this idea, as agents using private transportation, assumed to be more expensive than public transportation, are found to reside in the city centre while those using public transportation opt to live in the outer areas. A noteworthy observation is a shift in location preferences when the cost of public transportation drops significantly. In this scenario, agents using public transportation tend to reside in the city centre while those using private transportation move to the outer areas. This is a crucial finding that suggests that heavily subsidized public transportation can greatly influence the location choices of commuters. Currently, there is limited empirical research on the impact of public transport commuting costs on location choices, and as such, this finding from the model requires further validation.

The result of our final experiment shows that when agents are allowed to occupy larger plot areas, high income agents shift from the city intermediate zone to the city's outer areas, resulting in a linear increase in the average plot size with distance from the city centre. The desire for larger plots is one of the driving forces of suburbanization, as seen in many cities worldwide. A study by Kahn (2000) reported that cities in the USA, such as Chicago, Detroit, and New York, exhibit city-suburb land consumption differentials, with households occupying larger plots in the city’s outer areas compared to the city’s inner area. This is a crucial observation regarding the location choice of rich and poor households. By varying the maximum area, a household can occupy, a city can influence the location preferences of high and low income households.

6.6 Conclusion

In conclusion, the paper demonstrates that commuting and rental affordability play a crucial role in shaping the residential location choices of households, leading to a significant impact on urban morphology. Using an economic rational agent-based model of a hypothetical monocentric city, this study has made important observations that provide valuable policy directions for future research on the relationship between residential location choice, and affordability for different income groups. One of the main challenges in sustainable urbanization is to make cities more compact. Our model highlights the use of housing and transportation costs as a spatial policy tool to shape the urban form of a city. By increasing transportation costs and decreasing housing prices or rents, cities can promote compactness. At the same time, by optimizing rent and commuting prices, cities can become more affordable for residents and result in a more homogeneous density and income distribution.

Another issue in sustainable urbanization is income-wise segregation in cities. Our findings show that an increase in income inequality can lead to a more segregated city, with low income households getting confined to city inner areas having low commuting costs but high rent, thereby increasing their overall household expenditure as a percentage of their income. To address this, cities exhibiting such location patterns need intervention in housing prices to make housing affordable for low income groups residing near the city centre or job centres. Besides income inequality, our model also shows that plot size variability can contribute to the clustering of rich and poor households in cities, with rich households occupying larger plots in outer areas. Residential land policies that limit the maximum permissible plot area for households in a city can impact location choices and help reduce income-wise segregation.

The study has two important limitations. First, given that this study operates on the premise that agents are rational actors whose residential location choices are exclusively influenced by economic considerations, it would be prudent for subsequent investigations to incorporate the impact of non-economic and socio-cultural factors in shaping such decisions. Second, the model used in the study is designed for monocentric cities, which does not reflect the complexities of polycentric cities. Although monocentric urban forms can be seen in many cities around the world, the model can be improved to incorporate polycentric city designs.

Chapter 7

Simulating Built-up Expansion in West Delhi using a Neural Network Coupled Agent Based Prioritised Growth Model

Chapter Overview: The expansion of built-up areas is a complex phenomenon shaped by a range of spatial and aspatial factors that vary across space and time. Most of the previous studies have simulated land use patterns without considering the impact of futuristic development policies on land use. To address this gap, the study proposes a neural network coupled agent based prioritised growth model applied to the West region of Delhi. The model incorporates micro agents representing private developers who make land development decisions based on a cell's transition potential from non-built-up to built-up state, calculated by the neural network model. Macro agents, representing government planning agencies, enforce development constraints and provide incentives for development on a non-built-up cell through planned interventions. Simulations for 2021 demonstrate improved accuracy (kappa 0.85) with planned interventions compared to without any planned interventions (kappa 0.83), referred to as a business-as-usual scenario. The model also simulates land use for 2041 under these two scenarios. The resulting change in spatial growth under these two scenarios is visualised through a change map, which identifies areas of gain and loss in the built-up area as growth patterns shift from a business-as-usual scenario to a planned growth scenario. This model offers a useful tool for planners to understand where future growth is expected and how to channel the growth through strategic planning interventions.

7.1 Introduction

Urbanisation or built-up expansion is a complex issue that poses both opportunities and challenges for sustainable living, particularly in developing countries (Bikis, 2023; Abdulahi, 2022). To effectively manage urbanisation, modern planning aims at the incorporation of futuristic urban growth scenarios into urban development policies, with a focus on precisely identifying locations where built-up expansion is expected (Zhou et al., 2020). Predicting the spatial distribution of urban growth requires an understanding of the historical changes in urban expansion and modelling the rules that govern these changes.

Land use/land cover (LULC) analysis has been widely used and accepted as a method for analysing the changes in urban expansion over the last three to four decades (Gaur & Singh, 2023; Pan et al., 2022). In addition to monitoring urban growth, LULC analysis has been widely used to study different anthropogenic and ecological processes such as deforestation, forest fires, droughts, and floods. (Regasa et al., 2021; Kundu et al., 2017; Verburg et al., 2015). With the availability of high-resolution satellite imagery such as Landsat, Sentinel, MODIS, and image processing software such as GIS, ERDAS, and Google Earth Engine, LULC analysis can provide precise information about land use changes over time in a given region.

Modelling urban expansion can be traced back to the emergence of techniques such as Cellular Automata (CA) and Agent-Based Models (ABM), and the progress of Geographic Information Systems (GIS)-based software (Batty, 2005; Batty, 2007). The growth in CA models can be attributed to the work of Von Neumann and Standislav Ulam who developed simple CA models that could use local rules to generate mathematical patterns in 2-D and 3-D space (Ulam, 1976). Tobbler (1979) structured the CA models by defining the basic principles and essential components of a CA model. Conway's game of life showed the strength of CA models to mimic real-life situations by producing emergent behaviour that could not be predicted by the input data. With further refinements in CA, a fully operational CA model named SLEUTH was developed by Keith Clarke. Silva and Clarke (2002, 2005) showed the strength of the SLEUTH model to adapt to and thus, simulate different landscapes through model calibration, which made the SLEUTH model widely accepted by researchers globally and is still very much in use (Agyemang & Silva, 2019; Dadashpoor et al., 2019).

In the last two-three decades, developments in CA based urban simulation can be analysed from two aspects. First, while the conventional CA models were effective in modelling the spatial and temporal dynamics of land use change, they could not model the aspatial dynamics such as human behaviour, and their interactions and the decision-making process that plays an important role in determining the land use change. On the other hand, the agent-based (AB) models that emerged along with CA models, were effective in modelling the aspatial dynamics (Silva, 2011). The unique capabilities of CA and ABM led to the growth of a hybrid AB-CA approach in urban growth simulation.

The AB-CA models provided the necessary cellular approach to define spatial dynamics at the local level along with the agent-based approach to represent social interactions at the global level. While the traditional CA models were largely limited in their ability to model the global changes in the system, except for SLEUTH and CVCA models which could model the macro changes up to an extent, the hybrid models were proven to be more robust in accommodating the local and global changes in the system (Wu & Silva, 2010).

For this reason, the hybrid AB-CA models in the last few years have been used extensively to model the interaction of multiple agents in land/urban expansion. For example, Zhang et al (2015) used a multi-agent ABM to simulate the spatiotemporal change in land use in a coastal city in China. Saeedi (2018) built an ABM to simulate land use by factoring in the interaction between the micro-agents and the environment. Xu et al (2020) built a multi-agent ABM to simulate urban expansion in Auckland, New Zealand and found model accuracy better than single-agent ABMs.

The second important aspect in CA based models deals with their potential to frame transition rules which govern a cell's transition from one state to another. While in the conventional CA models, the transition rules were only governed by the state of neighbourhood cells, the hybrid CA models also take into account the impact of urban growth parameters on the cell transition (Xu et al., 2020). The transition rules can be user-specified mathematical functions or driven by a machine learning approach. Using a mathematical function to parameterise the influence of urban growth parameters on cell transition requires a model calibration approach to make the simulation realistic, for which studies use field data, expert analysis (such as analytical hierarchy process) and sensitivity analysis for different combinations of parameter values and their weights (Agyemang, 2022).

On the other hand, machine learning (ML) approaches automatically obtain the parameter values and their weights using training data and create rules of transition. Different machine learning algorithms have been used in previous studies to parametrise the model control factors such as Neural Networks, Decision Trees (Li et al., 2014), Random Forests (Kamusoko & Gamba, 2015), and Support Vector Machines (Ke et al., 2017), which are better able to capture the complex non-linearity in agents' transition rules (Islam et al., 2018).

In this context, artificial neural network, in the last two decades, has emerged as a more convenient and accurate method to model complex non-linear relationships between the urban growth driver factors and urban expansion. ANN models require fewer statistical assumptions as compared to regression models and thus, can be used with limited understanding of the relationships between the growth factors (Abiodun et al., 2018). Some studies also report higher accuracy in ANN models as compared to other machine learning methods such as random forest and support vector machines (Lazri & Ameur, 2018). Due to this many studies in recent years such as Liu et al. (2018), Zhao et al. (2019), and Xu et al. (2020) have coupled ANN with CA models to simulate land use.

While the previous studies have considered the aspatial dynamics and machine learning based transition rules in simulation, we find the urban growth parameters used in these studies to predict future land use change were static. Except for a few, studies have not simulated urban expansion considering urban growth parameters that emerge due to futuristic planned interventions. We find some studies in regional planning have considered the planning interventions to study future urban development. Liang (2018) simulated urban expansion in PDR, China incorporating future planning policies into a CA-based future land use simulation (FLUS) model. They found simulation accuracy improves as future planning policies are taken into account. A similar approach to model planning policies in CA based models can be seen in studies such as Chen et al. (2014) and Liu et al. (2017).

To improve the accuracy of urban growth simulation, it is essential to consider the impact of new infrastructure and residential projects planned by the government on land use change. In this study, we present a novel ANN-ABM based Prioritized Growth Model (PGM) with two main objectives. Firstly, we simulate land use for the year 2041 in the West Delhi district in Delhi, India, under the business as usual (BAU) scenario. Secondly, we incorporate a hypothetical scenario of expansion in metro rail services in the region to understand how this intervention will affect the spatial growth pattern in the year 2041 as compared to the BAU scenario. The contribution of this study lies in the incorporation of planning policies in ANN-ABM based growth models, demonstrating how future growth can be channelled to selected target locations. The PGM can also illustrate how growth can shift from one location to another when there is a deviation from the BAU scenario.

We assert that our model can be a valuable tool for urban planners to assess the potential effects of their proposed interventions on future growth patterns. Additionally, private land developers can benefit from the model's ability to predict how spatial growth in the region may vary if the government prioritizes growth in certain locations through planned interventions. The remainder of this paper is structured as follows. Section 7.2 provides the methodology used in the study. Section 7.3 provides the model results and discussion. Section 7.4 concludes the paper.

7.2 Methodology

7.2.1 Study area

The present study focuses on the Najafgarh tehsil, situated in the West Delhi district of Delhi, India. Bordered by the state of Haryana in the north, west, and south, the eastern boundary of the region is demarcated by the Najafgarh drain, separating it from the rest of Delhi. With a total area of 217.6 sq. km, the region comprises 1.31 sq. km of reserve forest and approximately 3.5 sq. km of the Najafgarh drain and water tanks/ponds. The 2011 census records a population of 332,720, of which 21.6% reside in the two urban towns and the remaining population in 48 villages, indicating a predominantly rural setting. The region has experienced a high population growth rate of 55.4% from 2001 to 2011 and 35% from 2011 to 2021, fuelled by increasing urbanisation in Delhi and the availability of affordable housing in the region. This has led to an influx of migrants from across the country, significantly impacting the region's land use pattern. The study region is shown in Figure 7.1.

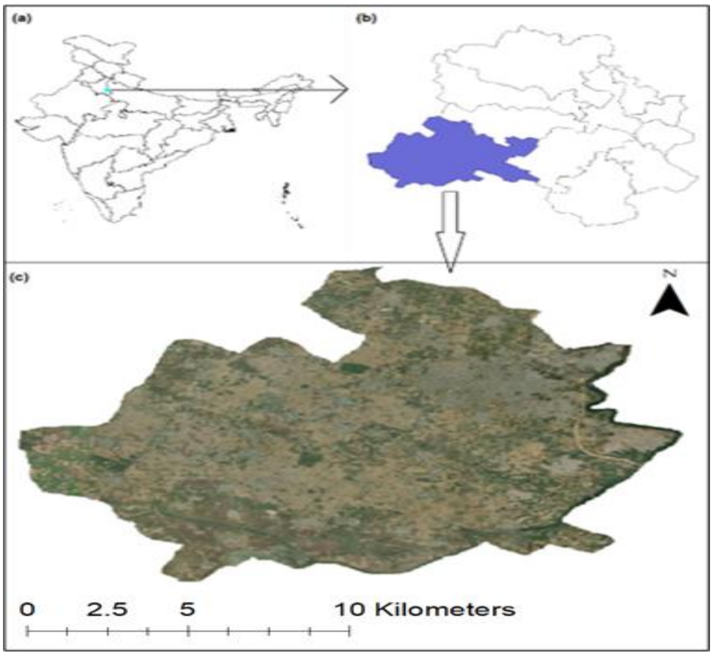

Figure 7.1: (a) Map of India showing Delhi (encircled). (b) Map of Delhi showing the West-Delhi district (blue coloured). (c) Satellite image of the studied region, Najafgarh tehsil lying inside the West-Delhi district.

The study's methodological framework is shown in Figure 7.2. It consists of three sequential steps – (1) Land use classification and change analysis, (2) Transition Potential Mapping, and (3) Agent Based Simulation.

7.2.2 Land use classification and change analysis

To assess changes in land use in the study area, we conducted a land use classification using satellite imagery from the Google Earth Engine (GEE) for the years 2001, 2011, and 2021. We utilized atmospherically corrected surface reflectance images from the Landsat Collection 2 database for the respective periods: Landsat 7 ETM+ for January 2001 to December 2001; Landsat 5 TM for January 2011 to December 2011; and Landsat 8 OLI/TIRS for January 2021 to December 2021.

The images were pre-processed using the CFMASK algorithm (GEE 2021) to remove noise and then classified using various supervised classification algorithms in the Google Earth Engine for each year. To validate the classified images, we used a sample database of 1,500 points obtained through stratified random sampling from Google Earth's historical imagery tool. Of these points, 70% were used for model training and the remaining 30% for model validation. We then computed the accuracy assessment by constructing a confusion matrix, which shows the sample of points that were incorrectly classified and is used to calculate the kappa index of agreement.

We utilized the Land Change Modeler (LCM) in the Idrisi Terrset software (Clark Labs, 2018) which simulates future land use changes through a three-step process comprising change analysis, transition potential modelling and change prediction.

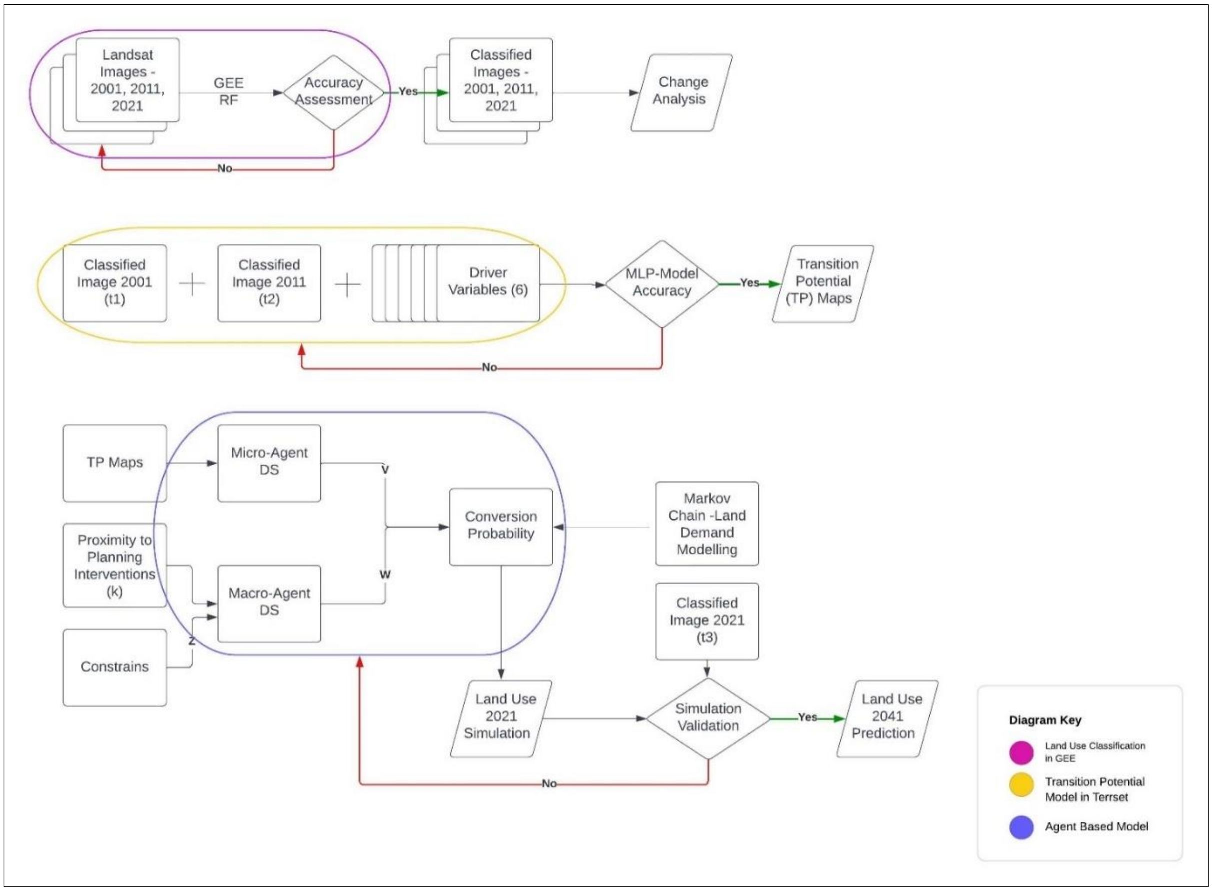

Figure 7.2: Study Methodological Framework

LCM has become a popular tool in recent years for studying land use change in various applications (Singh et al., 2022; Girma et al., 2022; Wang et al., 2021). The LCM is first used to analyse land use change for two time periods, 2001-2011 and 2011-2021. The model generated graphical representations of gains and losses in land use classes, as well as net changes experienced by each land class. After performing the change analysis, we ran the transition sub-model using the Multi-Layer Perceptron (MLP) neural network model.

7.2.3 Transition potential mapping using MLP

An MLP (multilayer perceptron) is a type of feedforward artificial neural network used to model complex non-linear relationships. It consists of an input layer, a hidden layer, and an output layer

(Eastman, 2016). In this study, the input layer comprises the land use image of 2001 and the urban growth driving factors, while the output layer includes the land use image of 2011. The MLP randomly selects cells that underwent land-use changes in the previous period. The training process uses 50% of the data and terminates after 10,000 repetitions and achieving the minimum RMS. The accuracy assessment of MLP is determined using the remaining half of the data, and it is computed using the expected accuracy and skill measure. The skill measure ranges from -1 to 1, with 1 indicating perfect forecasting, 0 indicating a random chance of predicting, and -1 indicating worse than chance (Girma et al., 2022; Gharaibeh et al., 2020).

$$S = \frac{A - E(A)}{1 - E(A)} \tag{7.1}$$

$$E(A) = \frac{1}{T + P} \tag{7.2}$$

where S is the skill measure, A is measured accuracy, E(A) is the excepted accuracy based on the number of transitions in the sub-model, T is the total transitions in the sub-model, and P is the permanency of classes in the sub-model.

The present study investigated various combinations of urban growth driving factors to identify the optimal combination for the MLP model and determined the relative impact of each variable on model accuracy through sensitivity analysis. Six variables were used in the model, including built-up density, distance to metro stations, distance to roads, distance to urban towns, distance to large rural settlements, and distance to small rural settlements. The density raster map was generated using the kernel density tool, and the distance raster maps were generated using the Euclidean distance tool in ArcGIS with vector data of the features. All input and output maps were processed to have the same spatial resolution (30m x 30m) and projection (WGS 84 UTM Zone 43N). After training and testing the model, a transition potential map was generated that indicates the likelihood of a cell transitioning from a non-built-up state to a built-up state in the future.

7.2.4 Land use simulation using MLP-MC-ABM

To simulate future land use, we utilized the Markov Chain (MC) model provided by LCM to calculate the change in demand for each land use category at a specified date. The MC model uses the change in land use between input and output maps to determine the amount of change that will occur in the future. First, the model builds a transition probability matrix that determines the likelihood of conversion between land use categories. Then, a transition area matrix is generated to show the change in total area (in cells) for each land use category over a specified number of years.

To account for development constraints and future planning scenarios in the simulated land use image, we have integrated the MLP-MC model with an agent-based model (ABM). The study region has been divided into grid cells of size 30x30 meters, referred to as patches, to execute the ABM. The future state of a patch is determined by the combined decision scores of the micro and macro agents. Private land developers serve as micro agents who choose a patch for development based on its transition potential calculated by the MLP model.

Macro agents refer to the government planning agencies responsible for setting development constraints and providing incentives for development in selected locations. Development constraints are rules that prohibit private land developers from building in certain areas. Incentives for development can be provided through the construction of transportation infrastructure, affordable housing, or job centres, among other things. This approach increases the likelihood of development in targeted areas. To analyse the impact of macro-agents on cell transition, we calculate the proximity of cells to areas where growth is prioritized. The final conversion probability of a cell from an urbanisable state to a built-up state depends on the nature and intensity of interaction between the agents, as described in Equations (7.3) to (7.8).

$$P_{t_1 \to t_2} = \left(V * DS^{micro} + \sum_{k=1}^{K} W_k * DS_k^{macro}\right) * S_{t_1} * Z \tag{7.3}$$

$$DS^{micro} = f_{ANN}(x_1, x_2, \ldots\ldots, x_n) \tag{7.4}$$

$$DS_k^{macro} = 1 - \frac{d_k - d_k^{min}}{d_k^{max} - d_k^{min}} \tag{7.5}$$

$$V + \sum_{k=1}^{K} W_k = 1 \tag{7.6}$$

$$Z = \begin{cases} 1, & outside\ the\ prohibited\ zone \\ 0, & inside\ the\ prohhibited\ zone \end{cases} \tag{7.7}$$

$$S_{t_1} = \begin{cases} 0, & cells\ already\ built-up \\ 1, & non-built-up\ cells \end{cases} \tag{7.8}$$

$P_{t_1 \to t_2}$ refers to the conversion probability of a cell from a non-built-up state to a built-up state during the period t_1 to t_2. DS^{micro} refers to the decision score of the micro-agents for a cell based on the cell's transition potential derived from urban growth driving factors, x_1 to x_n using the MLP-Neural Network model. DS_k^{macro} refers to the decision score of the macro-agent for a cell based on the cell's distance to the prioritized growth area k, denoted as d_k . d_k^{max} and d_k^{min} refers to the maximum and minimum distance from the prioritized growth area k across different cells, respectively. V and W are the weights assigned to the decision score of the micro and macro agents, respectively. The model allows for considering more than one prioritised growth area. The macro-agents also decide the development constraints, as denoted by Z. S_{t_1} denotes the state of the cell during the period t_1.

The cells are sorted in descending order based on their conversion probability scores. Then, cells with the highest conversion probability scores are selected according to the demand for cells in the simulation year t_2. This way, the MLP-MC-ABM model links past developments and future planning interventions, enabling the prediction of a cell's probability of conversion from a non-built-up to a built-up state.

We apply the aforementioned procedure to simulate land use for the years 2021 and 2041. The simulated land use of 2021 (comparison map) is validated using the classified Landsat image of 2021 (reference map). To accomplish this, we employ the VALIDATE module within the Terrset software, which uses the Kappa index statistics to assess the level of agreement between the two maps in terms of the quantity and location of cells in each land use category. A high validation score indicates the reliability and acceptance of the MLP-MC-ABM model in simulating future land use images.

7.3 Results

7.3.1 Land use classification and change analysis

Land use classification images were generated for 2001, 2011, and 2021, and consisted of four land use categories: built-up, cultivable land, barren soil/fallow land, and water, as depicted in Figure 7.3.

The classified images were validated using the historical imagery tool in Google Earth software and building the confusion matrix. Using the kappa coefficient and overall accuracy, the validation accuracy for the classified images was found to be highest using the random forest (RF) classification algorithm, as shown in Table 7.1. The study's land use classification accuracy falls within the acceptable range of greater than 80%, indicating a strong level of acceptance. Furthermore, this accuracy level is comparable to other recent studies that have used RF classifiers to classify land use, such as those conducted by Girma et al. (2022), Aslani et al. (2022), Becker et al. (2021), and Wang et al. (2021). Land use change was analysed during two-time intervals: 2001-2011 and 2011-2021. Table 7.2 lists the temporal changes in land use during these intervals. Figure 7.4 displays the gains and losses (in sq. km) and net change (in %) in the land use categories between the two-time intervals.

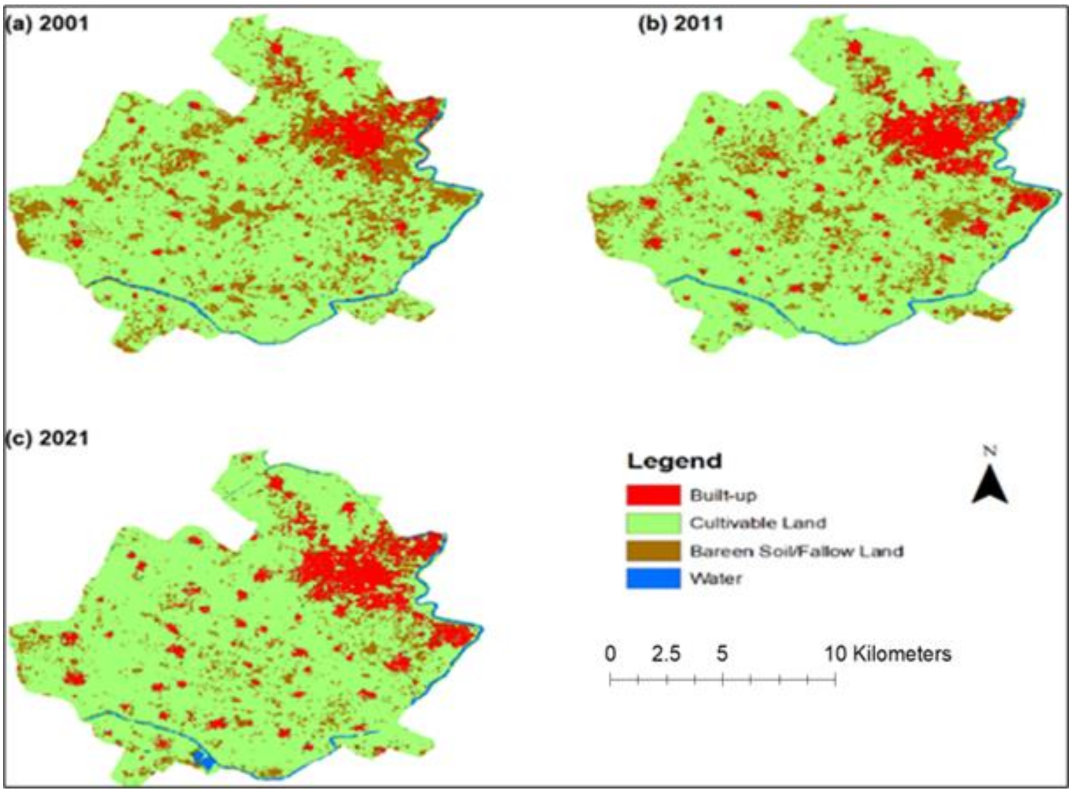

Figure 7.3: Classified Land Use maps of the studied region for years (a) 2001, (b) 2011, and (c) 2021 using GEE and RF algorithm.

Table 7.1: Overall Accuracy (OA) and Kappa coefficient for classified images

Year	2001	2011	2021
OA	0.87	0.88	0.92
Kappa	0.85	0.85	0.87

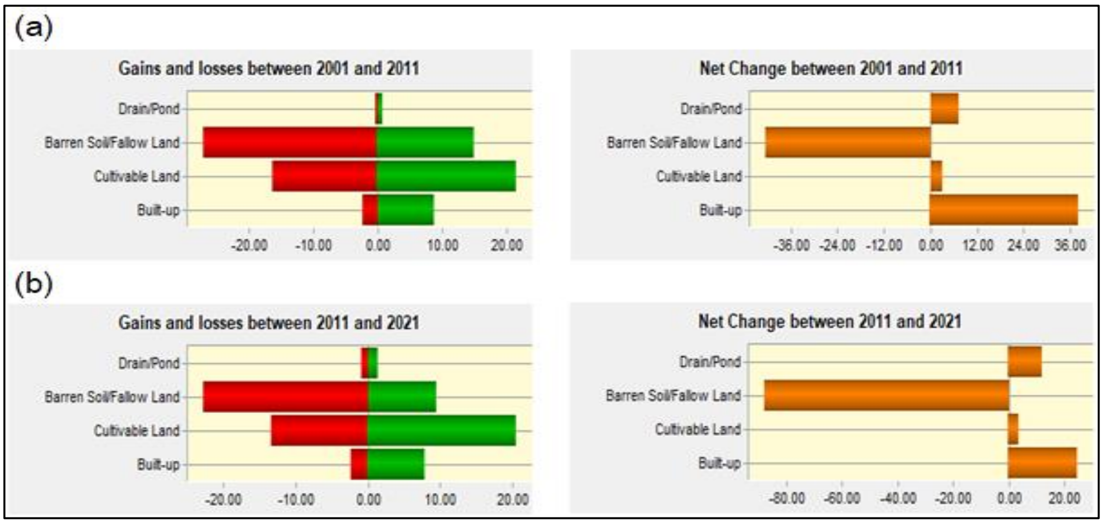

Figure 7.4: Gain and loss in land use (km2) and net change in area by land use category (in %) during the two-time intervals – (a) 2001-2011 and (b) 2011-2021

Table 7.2: Land use change analysis

	2001	2011	2021	Change 2001-2011	Change 2011-2021
Land Use	**Area (sq. km)**	**Area (sq. km)**	**Area (sq. km)**	**Area (%)**	**Area (%)**
Built-up	10.41	16.85	22.44	61.84	33.15
Cultivable Land	164.24	169.51	176.66	3.21	4.22
Barren Land	40.11	28.17	15	-29.78	-46.76
Water	2.87	3.1	3.54	7.82	14.11

Among the different land use categories, the built-up category witnessed the maximum percentage increase in the area during both time intervals. The built-up area grew by 62% during 2001-11 and 33% during 2011-21. Although there was a decline in the growth of built-up area during 2011-2021 as compared to 2001-2011, it was found to be higher when compared with Delhi's growth in built-up area of 24% during 2011-2021, as computed in different studies (Salem et al., 2021). The increase in the built-up area occurred at the expense of a 30% and 47% decline in the barren soil area during 2001-2011 and 2011-2021, respectively. The area under cultivable land and water also saw a considerable increase in both time intervals. To analyse the spatial pattern of land use change, we utilized the change map tool in the land change modeller (LCM) in Terrset. As depicted in Figure 7.5, during 2001-2011, the transition to the built-up area was primarily concentrated near the urban town of Najafgarh and around locations close to metro stations.

The construction of the metro station in 2005 significantly increased accessibility to these areas, resulting in their development. For instance, Qutubpur, a rural settlement located in the northwestern part of the region and near the metro station, experienced an eight-fold increase in population according to the 2011 Census. Similarly, areas such as Dindarpur and Goyla Khurd, also located near the metro station, saw a population increase of over 100%. In contrast, during 2011-2021, the transition to the built-up area was less intense and more evenly distributed across different locations, indicating that pre-existing built-up density played a significant role in driving growth during this period.

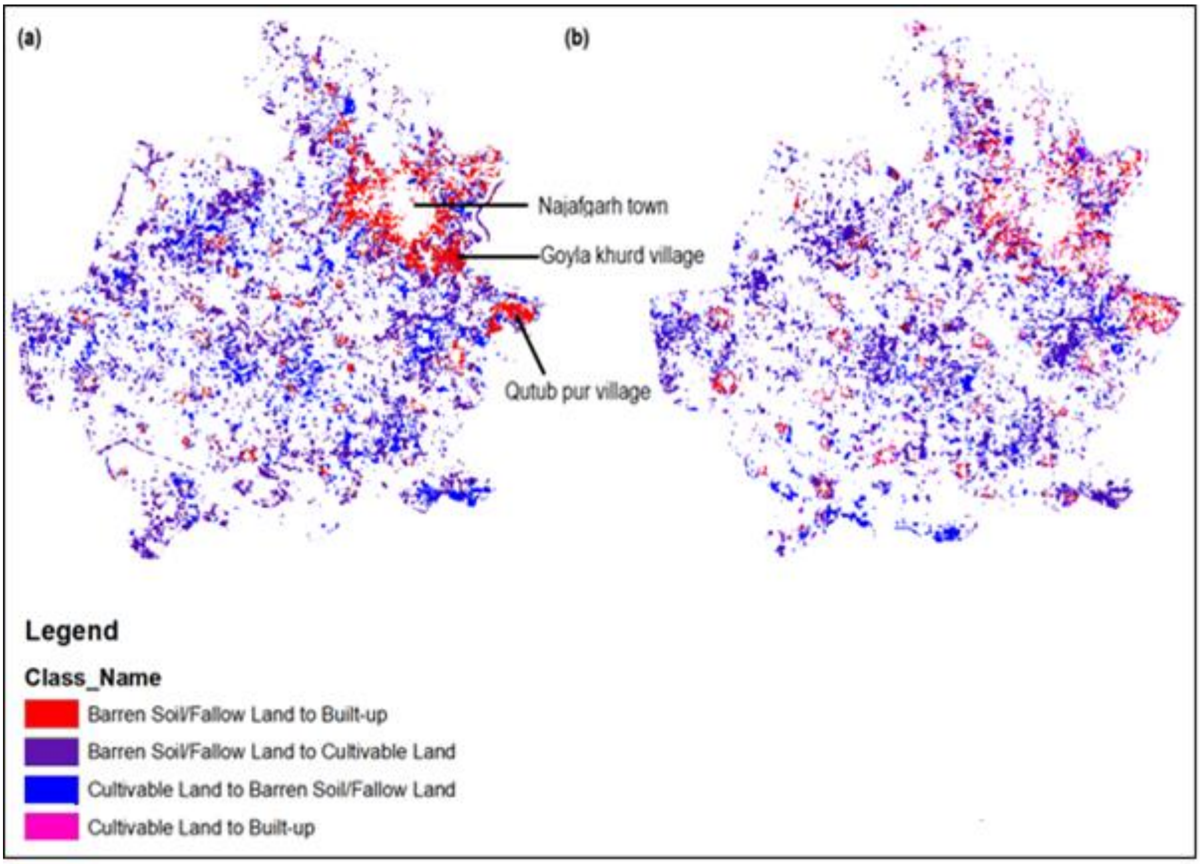

Figure 7.5: Change in land use between – (a) 2001 to 2011 and (b) 2011 to 2021

7.3.2 Variable selection and sensitivity analysis

The MLP model architecture comprises one input layer, one hidden layer, and one output layer. The input layer consists of 7 neurons, including the Land Use 2001 image and 6 urban growth driver variables, as shown in Figure 7.6. The model is trained using the land use 2011 image. The urban growth driver variables used in this study were selected based on their ability to drive urbanization, as reported in previous studies (Kim et al., 2020; Gharaibeh et al., 2020).

Built-up density is a critical factor in land use change, as vacant land near high-density settlements has a higher chance of being converted to built-up areas than land farther away. Areas with good accessibility to roads and metro stations are also preferred, as they facilitate commuting to workplaces. Proximity to urban towns and large villages provides multiple benefits to residents in terms of access to commercial and recreational services. We also include proximity to small villages as a driver variable, as these areas typically have lower land and rent prices, making them attractive to low-income migrants seeking to settle there.

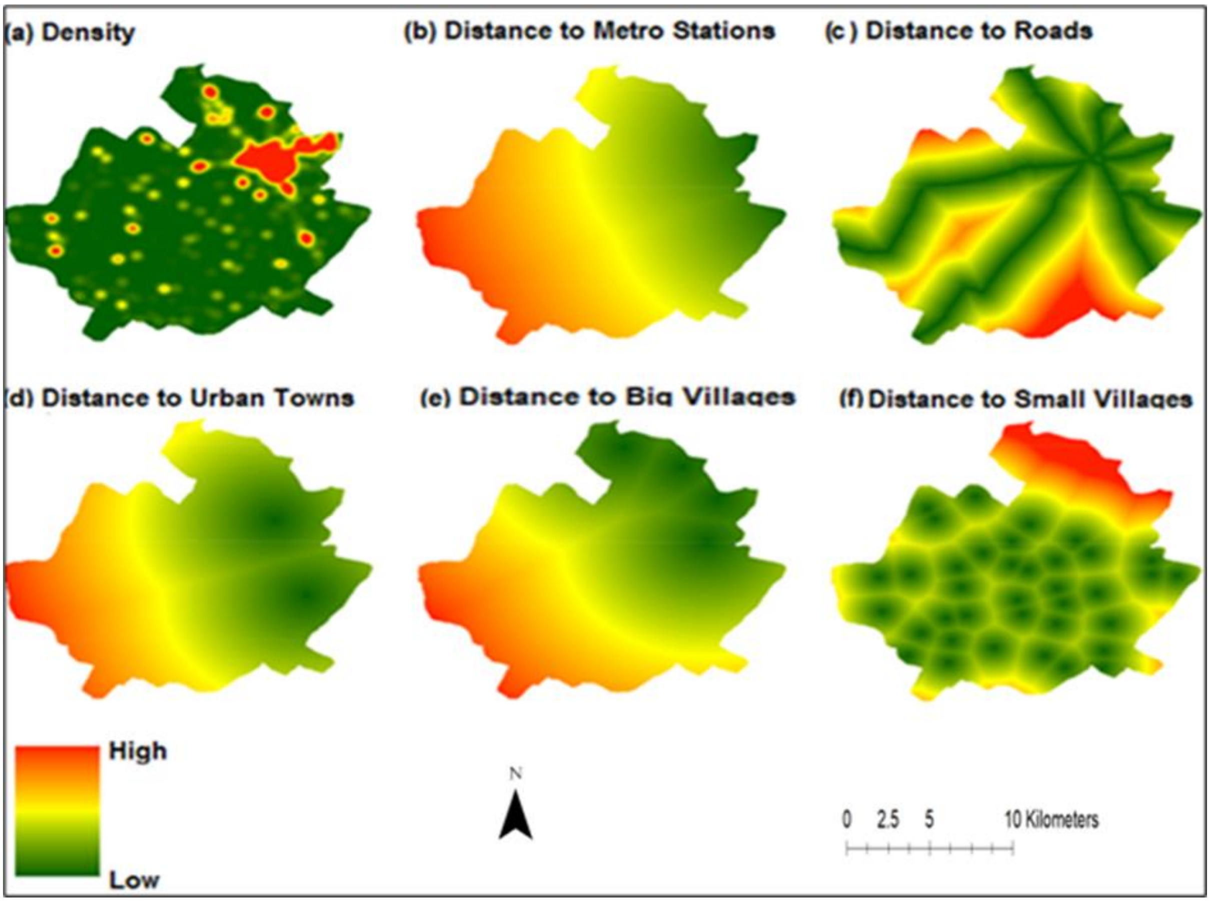

Figure 7.6: Urban growth driver variables used in the MLP model.

The MLP model's ability to simulate land use change was evaluated using the skill value and sensitivity analysis of the driver variables, which were computed after the model completed 10,000 iterations. The model demonstrated a high level of accuracy, achieving an overall accuracy of 92.21% and an overall skill value of 0.83, indicating its suitability for simulating future land use change. The model also provided sensitivity analysis for the variables, with the backwards stepwise constant forcing test showing which pair of variables had the least impact on model accuracy when held constant in a stepwise manner. However, the test results in Table 7.3 indicated that all variables had a significant impact on the model's accuracy, thus validating their inclusion in the model.

Table 7.3: Sensitivity analysis using backward stepwise constant forcing

Model	Variables included	Accuracy (%)	Skill measure
With all variables	All variables	92.21	0.8303
Step 1: var.[2] constant	[1,3,4,5,6]	89.15	0.8130
Step 2: var.[2,3] constant	[1,4,5,6]	86.14	0.8008
Step 3: var.[2,3,5] constant	[1,4,6]	75.32	0.7864
Step 4: var.[2,3,5,6] constant	[1,4]	74.94	0.5688
Step 5: var.[2,3,5,6,1] constant	[4]	71.02	0.3103
Variable 1 – Big Villages, Variable 2 – Urban towns, Variable 3 – Density, Variable 4 – Small villages, Variable 5 – Metro stations, Variable 6 – Roads			

The MLP model produces a transition potential map that depicts the likelihood of a cell transitioning to a built-up state. Cells with higher potential are more likely to undergo a transition. To simulate the land use for 2021, the model needs to determine the number of cells that will transition. The LCM includes a Markov chain (MC) process that creates a transition probability matrix (Table 7.4) that records the probability of each land use category transitioning to another. Afterwards, an area matrix is computed by multiplying the transition probability matrix with the cell count of each land use category. This provides the number of cells required to transition to a built-up state.

Table 7.4: Transition probability matrix from 2011-2021

Land Use	Built-up	Cultivable Land	Barren Land	Water
Built-up	0.9201	0.0642	0.0086	0.0071
Cultivable Land	0.0085	0.9005	0.0878	0.0032
Barren Land	0.1853	0.4889	0.3246	0.0012
Water	0.0007	0.1416	0.0042	0.8535

7.3.3 Simulation land use 2021 and validation

To simulate the land use in 2021, we use the MLP-MC-ABM using a set of Equations (7.3) to (7.8). The model allows for integrating the actions of micro and macro agents in the simulation. The DS^{micro} is the transition potential of a cell as calculated in the MLP model. One planning intervention by the government during 2011-2021 was the opening of two metro stations in "Najafgarh" town in 2019. We hypothesise that this would have incentivised growth in the nearby region. To accommodate the growth occurring due to the planned intervention, we use the macro agent's decision score for a cell, DS^{macro} based on the proximity of the cell to the nearest metro station (Eq. 7.5). Since the model has included only one planned intervention, the value of k is set to one. We also generate the constraints map where development cannot happen. This includes the area along the Najafgarh drain, forest area, vacant land lying under the central police force academy, and vacant land lying under the industrial area.

The combined effect of micro and macro agents is modelled in the ABM framework. The model is calibrated with different combinations of weights (V and W) assigned to DS^{macro} and DS^{micro}. In each case, the simulated image is validated with the actual Landsat classified image using the VALIDATE function in Terrset. The validation accuracy is measured using the kappa index of agreement, $k_{standard}$ between the two images. The variation in $k_{standard}$ with different values of weight (W) of DS^{macro} is plotted in Figure 7.7. The image with W = 0.1 provides the best validation accuracy and thus, is chosen to simulate the land use of 2021. The simulated image is shown in Figure 7.8 along with the actual land use 2021.

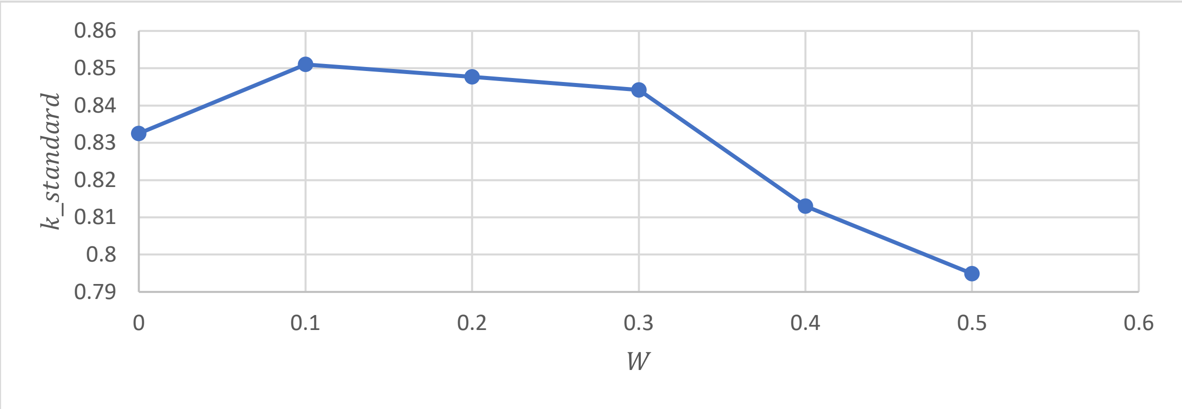

Figure 7.7: Variation in validation accuracy, $k_{standard}$ with weight, W assigned to the decision score of the macro-agent.

The inclusion of the planned intervention in the simulation of land use 2021 improves the accuracy of the model compared to a simulation without it. This highlights the benefits of incorporating futuristic planning interventions in our model. However, the low weight assigned to the decision score of macro agents suggests that the metro station has not yet attracted significant growth in its vicinity during the period of 2019-2021. This could be due to the metro station being operational for only two years and may require more time to stimulate development around it. Another factor could be the location of the metro station in an urban town with high built-up density, leaving less scope for further development.

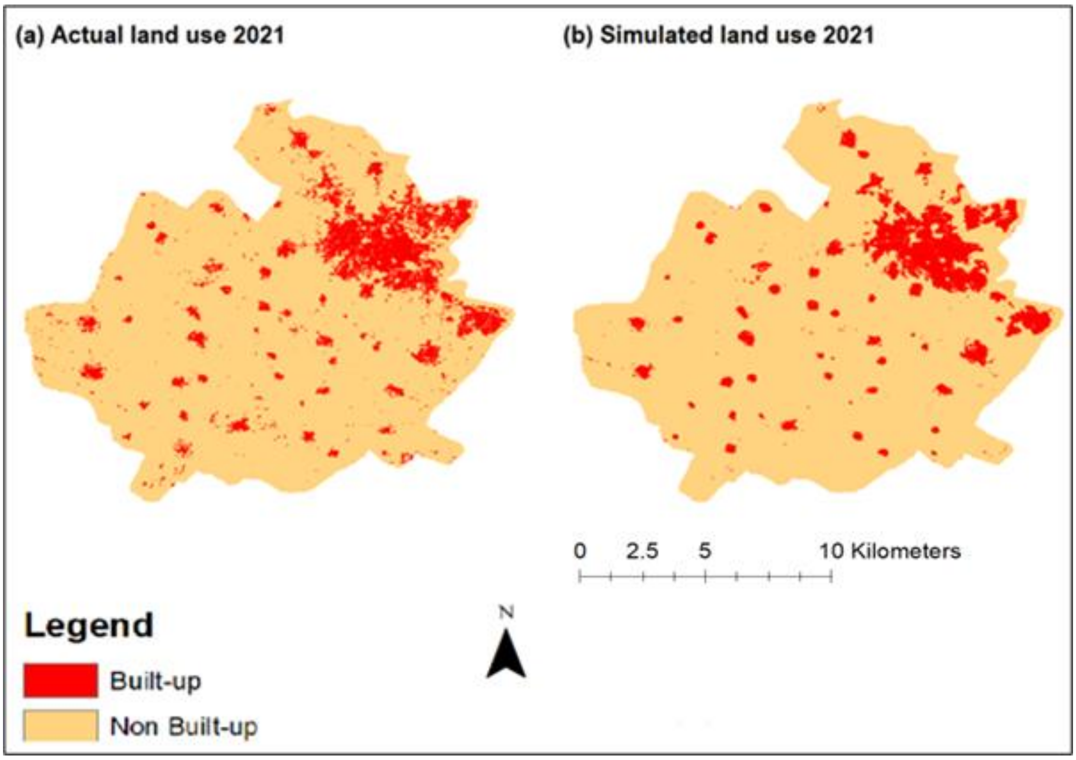

Figure 7.8: (a) Actual land use and (b) Simulated land use 2021

Validation result

The validation function calculates four different kappa indices, where a kappa value of 0% indicates that the level of agreement is equal to the agreement due to chance, and 100% indicates perfect agreement between the compared and referenced images. The traditional kappa index of agreement, denoted by $k_{standard}$, is 0.851. The overall agreement is shown by k_{no}, which is 0.86. $k_{location}$ is equal to 0.89, indicating the extent to which the two maps agree in terms of the location of each category. Lastly, $k_{quantity}$ is equal to 0.87, indicating the extent to which the two maps agree in terms of the quantity of each category. All the kappa indices are well above the satisfactory range (>80%) which denotes high reliability in our model in simulating future land use. The VALIDATE function also generates a bar graph that shows the overall proportion of cells correctly classified, which in our model is 90.4%, as shown in Figure 7.9.

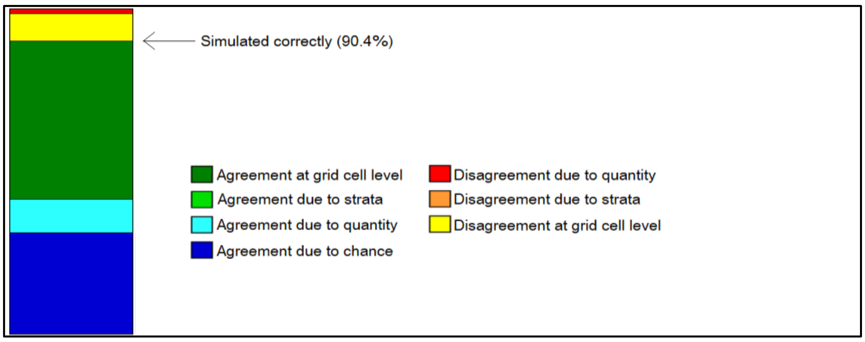

Figure 7.9: Simulation accuracy

7.3.4 Simulating land use 2041

One of the objectives of the study is to simulate the spatial growth in built-up areas in 2041. The model after being validated successfully is first used to generate the land use of 2041 under the assumption that there is no planned intervention by the government during 2021-2041 which means the weight assigned to the decision score of the macro agent is nil. This is what we refer to as the business-as-usual scenario. The predicted image of 2041 is shown in Figure 7.10 (a). The built-up area in 2041 in the region is projected to be 32 sq. km registering a growth of 43% from 2021 to 2041. The growth is anticipated to occur on the periphery of existing urban and rural settlements rather than being concentrated solely near the urban towns.

Another objective of this study is to analyse the spatial variation in future growth when development at a location is prioritized through infrastructure planning. One of the infrastructure planning interventions by the government since 2003 across Delhi has been the construction of metro rail. Based on our understanding of the future growth of metro rail services in Delhi, we hypothesize the construction of 2 metro rail stations in the study region at locations marked in Figure 7.11.

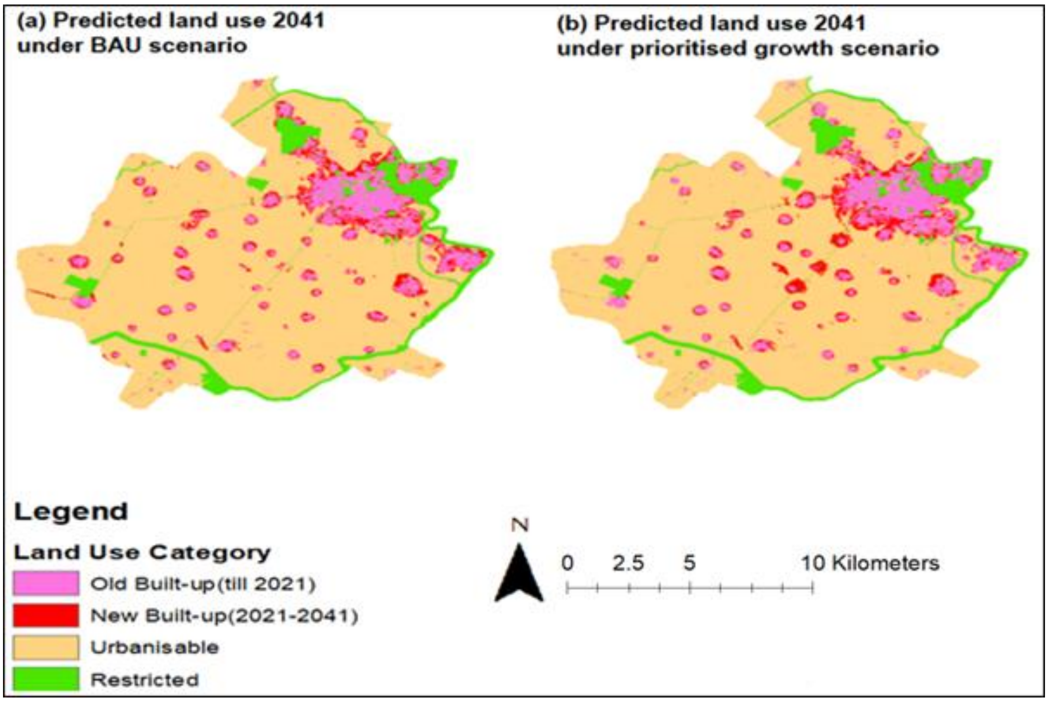

Figure 7.10: Predicted land use 2041 under (a) BAU scenario and (b) Prioritised growth scenario

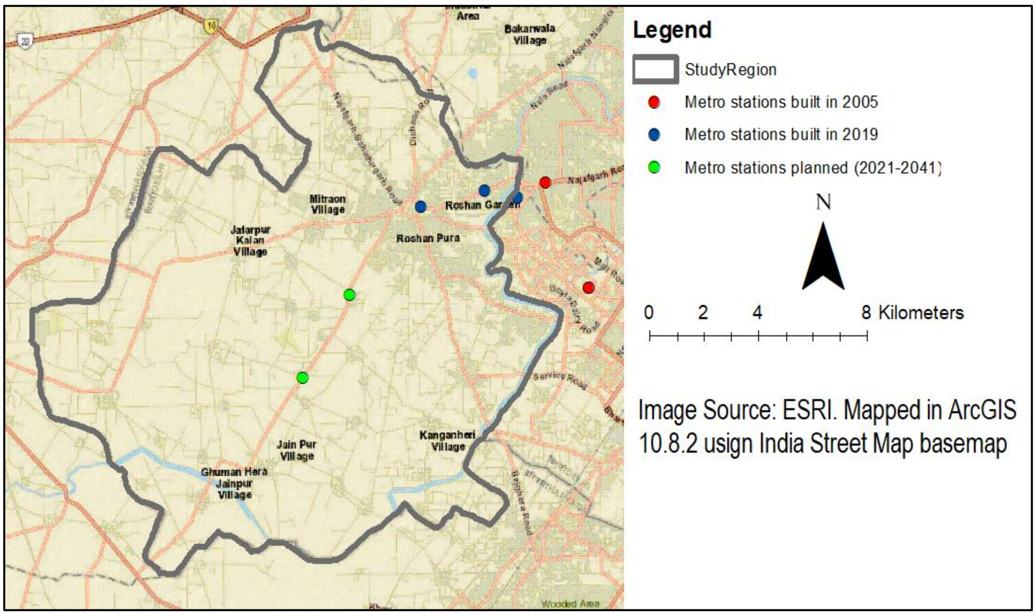

Figure 7.11: Map showing the location of assumed upcoming metro stations in the region.

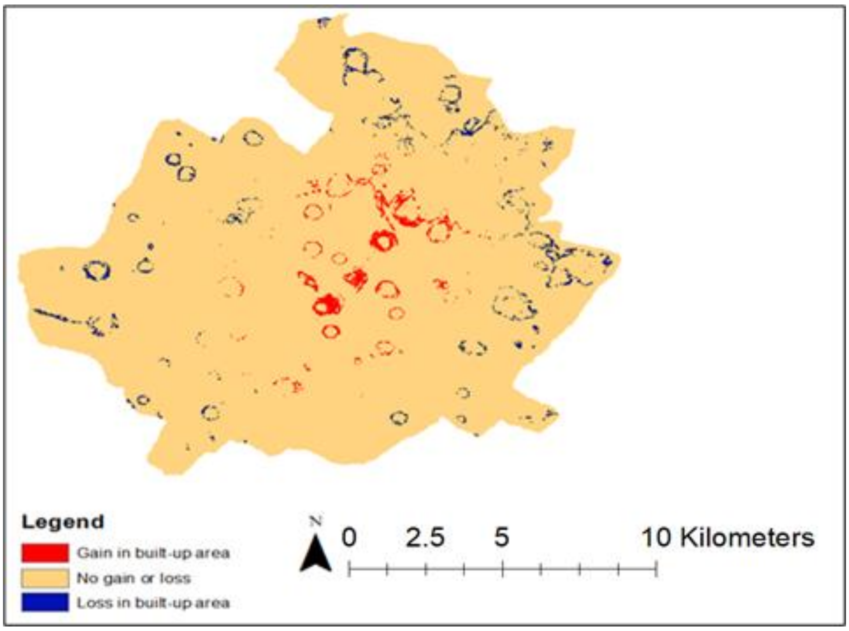

Figure 7.12: Change in spatial growth during 2021-2041

The ANN-ABM framework enables us to visualize how the growth in built-up areas by 2041 can shift from the business-as-usual (BAU) scenario following the opening of two metro stations. This agent-based modelling framework combines growth projections based on BAU predictions with those derived from futuristic planning interventions.

In this framework, micro agents are private developers who settle on vacant plots that they believe have a higher probability of built-up, while macro agents aim to develop plots close to the new metro stations. Taking into account historical development and government-prioritized growth areas, the probability of a cell transitioning to a built-up category is calculated. We assume that both agents influence each other equally and assign equal weights to their decision scores. Figure 7.10(b) illustrates the predicted land use in 2041 under the planned intervention.

We created a change map to highlight the areas where growth is expected to shift from the business-as-usual scenario to the planned growth scenario. In Figure 7.12, the red areas represent locations that will gain built-up area, while the blue areas represent locations that will lose built-up area as development around metro stations is prioritized. As a result of this infrastructure planning, settlements on the periphery of the region are expected to experience lower growth in built-up areas than they would have otherwise.

This demonstrates the usefulness of the MLP-ABM in modelling the dynamics of future growth based on planning interventions. The studied region has a high potential for growth in the coming decades, as the majority of the land is either barren or under cultivation. As the rate of urbanization in Delhi increases, there will be a significant influx of migrants to this region, making planning interventions critical for sustainable urbanization. Since the region comprises both low-density rural and high-density urban settlements, different planning interventions can be implemented to make the future urban form sustainable. For example, developing transportation, economic, and recreational services around rural areas or building low-cost residential flats or parks/green areas near high-density areas.

Regardless of the type of planning intervention, our model is capable of illustrating how the spatial pattern of growth will shift from the business-as-usual scenario to a planned growth scenario. Therefore, this model can aid planners in predicting and directing future growth while also helping private developers identify areas that are subject to change and thus vulnerable to growth based on planned interventions.

7.4 Conclusion

Understanding where future development will occur can assist city planners in efficiently managing urbanization and associated challenges. While previous studies have been able to simulate existing land use through the coupling of ANN-ABM, the impact of futuristic development policies on land use simulation has largely been ignored. This can lead to an inaccurate assessment of land use and undermine the role of planning agencies in determining future land use patterns.

To overcome this issue, this paper proposes an integrated ANN-ABM based Prioritized Growth Model that considers the roles of micro and macro agents in the region's development. The micro agents prioritize growth based on historical trends, which are computed using an MLP-Neural Network model, while macro agents prioritize growth based on futuristic planned interventions. The final conversion probability is derived based on the intensity of interaction between the micro and macro agents.

The Prioritised Growth Model operates sequentially, beginning with land use classification using the Google Earth Engine (GEE) platform. The RF classifier was employed to classify land satellite images of 2001, 2011 and 2021, which yielded a high degree of classification accuracy. Land use changes during 2001-2011 and 2011-2021 were computed and visualized using Terrset software, indicating a 62% and 33% increase in a built-up area in the region during these periods, respectively. The model was then applied to simulate future land use in the West Delhi region of Delhi, India. The MLP model in Terrset was used to build a transition potential map for 2021 based on the land use map of 2001 and 2011, along with six urban growth driver variables.

The simulated land use in 2021 with the impact of planned interventions showed higher accuracy (kappa 0.85) compared to the simulation without planned interventions (kappa 0.83). The model simulates land use in 2041 under the business-as-usual scenario and under the prioritised growth scenario, which assumes futuristic growth along the two newly constructed metro stations. A change map is used to visualize the shift in spatial growth from the business-as-usual scenario to the planned growth scenario, showing locations with gains and losses in built-up areas.

The ANN-ABM prioritized growth model proposed in this study highlights the potential of planning interventions to shape future growth patterns. By providing insights into where future growth is expected to occur and how it can be channelled to prioritized locations through suitable planning interventions, the model can be a useful tool for urban planners. This approach can lead to more effective measurement and evaluation of planning interventions and help ensure that growth is managed sustainably.

Chapter 8

Discussion and Conclusion

8.1 Revisiting Study Research Framework

The study aimed to examine the dynamics of urban form, flows and accessibility, in the city of Delhi, utilizing geo-computational methods. Characterising cities as complex adaptive systems, we built a conceptual framework using three crucial components of urban science which are, space, flow and human behaviour. We apply this framework to the city of Delhi to examine some of the important phenomena which are fundamental to urban living.

The issues examined in this study dealt with – sustainable commuting, equitable distribution of services, sustainable built up-forms, city affordability and residential location choice, and simulating the built-up expansion. These issues were examined in the thesis using the following questions:

(a) What factors determine the individual's choice of commuting distance and travel mode to the workplace and up to what extent do the cities and neighbourhood's built environment influence the commuting behaviour?

(b) How does travel attitude and residential location choice influence the causal linkage between the built environment and commuting behaviour?

(c) Is the spatial distribution of services across different neighbourhoods in Delhi equitable or biased towards neighbourhoods of a particular socio-economic characteristic?

(d) How can one analyse the heterogeneity in the urban form of a city? Does urbanisation in cities like Delhi is resulting in unsustainable living when analysed through the elements of urban form?

(e) How does city affordability impact the residential location choice of households and result in different density profiles? What economic policies can be sought that make cities optimally dense and more affordable for low-income groups?

(f) How can the future expansion in the built-up area be mapped and analysed using simulation techniques and how does the spatial growth vary taking into account the impact of futuristic planning interventions?

To examine these issues, the study utilised different datasets and methodological approaches. For empirical work in Chapter 3, we collected primary data through a field survey in Delhi where we interviewed 1679 working individuals about their commuting behaviour, socio-economic characteristics, neighbourhood characteristics, and travel attitude. The survey findings were analysed using descriptive statistics and regression methods.

In the rest of the empirical chapters, we used secondary data from different sources. In Chapter 4, to examine neighbourhood accessibility, we used location data of different services publicly available online on the government department website and geocoded these locations in GIS software. We also created neighbourhood maps and assigned the neighbourhoods with their population score using the neighbourhood population database from the state election commission 2022. In Chapter 5, to examine variation in urban form, we extracted elements of urban design by processing land satellite imageries, and the open street map database. We used Google Earth software to map residential areas in Delhi and used a grid-based technique to draw neighbourhood boundaries. In Chapter 7, to model the built-up expansion in Delhi, we created a land use raster dataset of different periods through land use classification. We also used different spatial and statistical datasets from Delhi master plans, Delhi municipal corporation, and Indian census documents for the study region.

We now provide a summary of key findings from the study.

8.2 Summarising Study Key Findings

Chapter 3 aimed at understanding the relationship between commuting behaviour and the built environment. Previously, not many studies have examined such linkage for cities in the global south primarily due to the unavailability of travel-related data. Moreover, only a few of them considered the influence of travel attitude on commuting behaviour. Ignorance of travel attitude, as some studies find, may lead to an overestimation of the impact of the built environment on commuting behaviour and result in less realistic policy recommendations. Overcoming these two concerns, our case study of Delhi was built on the household survey data incorporating the travel attitude of the commuters.

Our results analysis shows a pattern of outward commuting in Delhi, as those living in the city's inner area are more likely to commute longer distances to workplaces. We also find that the majority of the respondents have a preference to select a residential location near to workplace to minimise their workplace distance. Concerning mode choice, we find respondents whose workplaces lied in the city’s outer area are more likely to use a car to commute to the workplace. On the other hand, those who live in the city’s outer areas are more likely to use public transportation. While our results show that proximity to metro stations has an impact on mode choice, we also notice that it is the preference of commuters to use public transport which makes them select their residential location near transit stations. This means that, while the built environment appears to influence the mode choice, the causal mechanism is the travel attitude which may influence the choice of the built environment which then relates to the mode choice. Our results also show that people tend to rely on private vehicles over public transport due to poor connectivity to transit stations and to decrease commuting time. Overall the study finds that the influence of travel attitude expressed in terms of commuting cost, time and comfort is important for understanding commuting behaviour.

Chapter 4 examined the inequity in accessibility to services for different socio-economic neighbourhoods in the city of Delhi using a geographically weighted regression model. One of the challenges in this work was to create the spatial database of all residential areas in Delhi with their population size and socio-economic indicators. We found fewer studies have examined the spatial distribution of services across neighbourhoods in cities from the global south using the social equity perspective. On this ground, the study becomes significant as it measures the accessibility for every residential location in Delhi, making it the most comprehensive study, best to our knowledge, on accessibility to services in Delhi. The study finds that there exists spatial inequity in accessibility to services in Delhi. However, the inequity is more attributed to neighbourhood spatial location rather than to their socio-economic characteristics. Although neighbourhoods with low income and a high percentage of scheduled caste population were found to have low accessibility to some of the services, it could be said only for a few specific neighbourhoods' clusters and cannot be generalised for the entire city mainly due to the non-stationary relationship between the variables.

Chapter 5 explored the residential built-up forms typologies and assessed their impact on sustainable urbanisation in Delhi. The study of urban form in developing cities like Delhi needs to take into account the diversity of micro-scale urban form features, which studies in past have mostly failed to do so due to the lack of neighbourhood maps and spatial data. The study mapped and analysed the variation in the urban form at the neighbourhood level using a grid-based k-means clustering technique. The study finds the urban form in Delhi can be clustered into six typologies. These typologies were then characterised by their dominant urban form feature using a machine learning-based approach and examined for the degree of urban sustainability. The results show that only 19% of the residential built-up area in Delhi provides sustainable living when analysed from the urban form perspective. The study's methodological approach becomes important as it emphasises incorporating the built-up form in the characterisation and measurement of urbanisation. The study is also significant as it shows that, as cities like Delhi urbanise in the coming years, the emerging urban form will require careful planning interventions in the built-up design features to enhance sustainable living.

Chapter 6 examined the impact of variations in the commuting cost and housing rent on the individuals' residential location choice and the complex urban pattern that results from it. The study builds an agent based model as a proof of concept to understand how the density pattern changes with changes in the city affordability in a monocentric city. The simulation results highlight the use of housing and transportation costs as a spatial policy tool to shape the urban form of a city. By increasing transportation costs and decreasing housing prices or rents, cities can promote compactness. At the same time, by optimizing rent and commuting prices, cities can become more affordable for residents and result in a more homogeneous density and income distribution pattern. The study also shows that an increase in income inequality and land ownership can lead to income-based residential segregation. The study becomes significant as it shows the potential of agent based modelling approach in simulating the complex urban growth which evolves from the individual choice of residential locations.

Chapter 7 aimed at simulating the built-up expansion in the west Delhi region using the neural network coupled agent based model. While previous studies have used different simulation techniques to model built-up expansion, most of them have not considered the impact of futuristic development policies on land use. Considering this caveat, the study proposes a neural network coupled agent based prioritised growth model that simulates growth driven due to historic factors and futuristic planning interventions. The simulations for 2021 demonstrate improved accuracy (kappa 0.85) with planned interventions compared to without any planned intervention (kappa 0.83), highlighting the need to include growth driven by planned interventions in future simulations. The model then simulates land use for 2041 under two different scenarios. First, the model simulates land use in 2041 under the business-as-usual scenario without considering the futuristic planning interventions. Second, the simulation happens under the prioritised growth scenario, which assumes futuristic growth along the two hypothesised metro stations. A change map is drawn to visualize the shift in spatial growth from the business-as-usual scenario to the planned growth scenario, showing locations with gains and losses in built-up areas. The study becomes significant as the proposed model in this study can aid planners to predict and direct future growth and can also help private developers to identify areas that are subject to change and thus vulnerable to growth based on planned interventions.

8.3 Policy Recommendations: Towards Accessible and Sustainable Cities

Urban planning in the era of rapid urbanisation has become more important than ever. The future of cities depends on the planning interventions planners take today. Cities being a complex adaptive system cannot be predicted as they are always evolving but can surely be invented by understanding the dynamics that shape the city's form and functions. Inventing future cities becomes important because cities are not only spaces for economic growth, but they also provide opportunities to build social, economic, and cultural connections among different sections of society which is beneficial especially to those who are socially or economically marginalised. Thus, planning measures become important to invent future cities that enhance the quality of life for all its residents.

This study, after examining the crucial components and their interaction with the city of Delhi, proposes some policy interventions that we believe are essential to make cities like Delhi more sustainable. We believe behind every policy and planning intervention there is an underlying objective and value-laden vision that guides the interventions. The study proposed policy and planning interventions are guided by the following considerations –

1. Building cities that enhance accessibility to services and workplaces, and not just enhance mobility through building travel infrastructure.

2. Cities need to provide better accessibility to public transport with a twin goal to reduce vehicular congestion and carbon emissions.

3. Spatial distribution of services in the city should result in social justice.

4. Urbanisation in developing cities needs to be de-aligned with the notions of population or building density. For sustainable urban living, the measure of urbanisation should factor nature of the built-up form.

5. Residential location choice driven by economic factors can play an important role in changing the city density pattern. Concerns for city affordability for different income groups should be examined in planning the city development plans.

6. Predicting future built-up expansion using machine learning based simulation should also factor human interventions and not just be solely driven by past growth patterns.

Based on these considerations we now discuss key policy measures which can be implemented to make Delhi an accessible and sustainable urban form.

1. Building local employment clusters: Based on our findings, the study advocates for making Delhi a megapolis that will require a dense and interconnected network of multiple employment clusters connected with a fast travel infrastructure. As noted in the study results, the city currently has two employment clusters, they fall in the city's outer areas which makes commuting time intensive and has adverse impacts on environment sustainability. Few industrial clusters are present in the city's central areas, but they do not provide employment in emerging tech sectors and thus are not seen as an employment hub for the youth. In this context, we see that Delhi require multiple employment clusters spread across different parts of the city. Rather than planning for a giant employment zone in a city, the focus should be laid on how local employment clusters can be created that can act as a node of employment generation and can increase accessibility to the workplace. These new clusters should

focus on providing jobs in the area of ICT and business consulting which are in demand and have the potential to enhance the city's economy.

To create multiple employment clusters in Delhi, planning needs to focus on a few important aspects such as (a) Infrastructure investments: The allocation of infrastructure projects may be influenced by political considerations, favouring certain areas or constituencies over others. This can hinder the equitable development of employment clusters. (b) Transportation: Inadequate transportation infrastructure, including roads, public transit, and last-mile connectivity, can limit the accessibility of potential job clusters and hinder the ease of commuting. (c) Land Scarcity: Delhi's limited available land can make it challenging to allocate sufficient space for multiple employment clusters, especially when there is competition for land for other purposes like housing. (d) Market Forces: Economic factors, including market demand and industry trends, can influence the feasibility of establishing employment clusters in specific areas.

The above highlighted challenges are important to consider while planning job clusters in the city. Examples from many developing countries show that with creation of public infrastructure like transit stations or job clusters, leads to the gentrification of the area. By which the local population, usually belonging to low-income households gets displaced owing to political actions like slum clearance policy or due to market forces like an increase in housing rent. Thus, the policy of multiple employment clusters should incorporate provisions like in-situ development of slums and additional houses for low-income groups near the site of planned job clusters.

2. Regulatory measures to disincentivise car usage: Our findings suggest that the use of cars in Delhi is more attributable to the household's socio-economic and travel attributes than to built-environment features. As Delhi urbanises and continue to advance its per capita economic growth and road side infrastructure, the use of car is expected to increase. In such scenarios, policy makers should not only rely on built environment measures such as transit-oriented development but should also focus on economic measures to discourage the use of cars in daily commuting to the workplace. One such measure can be to impose congestion pricing or road tax during peak hours. Similarly, policies like green tax credits for individuals who use public transportation, carpool, or purchase electric, or hybrid vehicles should be implemented. Also, policies like the even-odd car use schemes should be regularly implemented to allow people to explore alternative means of commuting. Some reward programs can be implemented such as free vouchers to users of public transport, free travel to women travellers and senior citizens in public transportation on some days in a week.

We also find in our study that travel attitude related to comfort maximisation makes people use private travel mode. Measures to make public transportation more comfortable, Mobility as a Service (MaaS) should be developed that can integrate various transportation options (public transit, ridesharing, bike-sharing, etc.) into a single, user-friendly system. Also, concern about safety in public transportation is important to address, especially for female travellers commuting during odd hours. The availability of female police force in buses, metro rails and transit stations is required to make female travellers feel safe.

Bringing change in people's preference for cars is a daunting task. It requires a change in cultural norms that delinks income or economic status with car usage. In the global south where a car is seen as a symbol of economic status, the choice of commuting by car is many a times a society-imposed norm. Urban planners should aim to break this linkage between income and car usage by creating more awareness about the larger benefits of using public transport. Involving community participation in the planning of public transport is also a right step that may enhance the sense of ownership among households and change their preference towards the use of public transport.

3. Urban planning for social equity: Our study findings advocate that the very practice of urban planning needs to shift its focus from infrastructure development to socio-economic development. Land use and transportation development are considered two major urban planning components. However, they do not directly result in inclusive development of the city. A third pillar of urban planning which has received limited attention especially in the global south is the planning for social equity. Resource or service planning in a city should be done keeping the community characteristics at the centre stage. Our study shows that some poor socio-economic neighbourhoods in Delhi have low access to services. Identifications of such neighbourhoods thus become important to precisely plan service location which requires the availability of neighbourhood maps and spatial data on neighbourhoods' mean socio-economic characteristics.

In the absence of such data, planning is done at a low level of spatial resolution such as wards or districts that do not showcase the spatial variation in access to services. As cities like Delhi have high income inequality and social diversity, planning for social equity can only happen with high-resolution spatial data. Thus, the study recommends the urban policy in Delhi should focus on creating neighbourhood maps and spatial datasets on household socio-economic characteristics so that planning can happen to achieve social equity in access to services.

4. Policy for urban space: Acknowledging the presence of informality in urban space, the paper argues for building urban policies that account for the spatial mobility in urban space in contrast to the norms of spatial fixity which is more evident in existing urban policies. The norm of spatial fixity assumes the urban is a homogenous entity across a city. As our results showcase, there exists heterogeneity in the urban form features across urban space that makes urban a spatial phenomenon and calls for invoking the understanding of spatiality in measuring the 'urban' within a city.

Our study also showcases how high-resolution geospatial data on an urban form can be used to map the urban form heterogeneity in Delhi and further link it to analyse the extent of sustainable urbanisation. Our results show how unplanned urbanisation can result in the creation of unsustainable living spaces. While unsustainable urban formations like slums and informal housing arise as unintended consequences of governmental actions, their widespread presence and resilience are deeply intertwined with socio-economic and political factors. Despite the absence of secure property rights and substandard living conditions, these neighbourhoods endure because they serve multiple roles. On one hand, they function as intricate economic and social systems, offering affordable housing and livelihood opportunities for marginalized labourers while contributing to the broader urban economy. Slums also play a vital part in the social support system for both the working and middle-class populations by providing affordable housing options. As Weinstein (2014, p. 27) suggests, the Indian government often adopts a policy of "supportive neglect" towards slums due to their cost-effective response to the pressing issue of housing shortages in cities. Additionally, a thriving informal real estate market has taken root within Mumbai's slums (Chattaraj 2016). It is the social utility of these slum areas that leads the state to tolerate their existence, resulting in their ongoing proliferation and persistence.

Given this complex interplay between form and functions, such structures are going to stay and rise in the city, as has been the case previously. Rather than planning for slum clearance or not regularising planning in such areas, planning measures should focus on enhancing the quality of life and generating a ‘sense of place’ by rooting the planning in the place's socio-economy, culture, and history. In a nutshell, Delhi urgently needs a holistic urban policy framework to manage the growing urbanisation and stop the proliferation of slums and illegal housing.

8.4 Concluding Remarks

Cities, as a complex system, can be studied under different theories, ideas, models and perspectives. Given the fact that cities are dynamic entities and are continuously evolving, they cannot be studied from a top-down approach. However, cities whether big or small, developed or developing, do have some similar basic constructs which can be studied to understand the nature of their growth pattern and future evolution.

The study developed a conceptual framework of spaces, flows and human behaviour which forms the fundamental basis of urban development, to study and examine the three key components – (a) accessibility to services, (b) built environment and travel behaviour, and (c) urban form and built-up expansion, and their interrelationships, in the city of Delhi. The study aimed to understand how the different physical and non-physical components of the city influence each other, and what planning measures can be taken to make the city form and future urbanisation environmentally sustainable and beneficial for the different socio-economic groups. In this endeavour, the study explored answers to some of the critical questions in urban science such as, why people commute longer distances to the workplace and what factors affect their choice of travel mode? How are different services distributed in a city and who gets access to these services? Does urbanisation result in the rise of unsustainable built-up forms? How does built-up expansion and density pattern in a city change with city affordability? How can urban planning be made more efficient by incorporating the planner's decisions in simulation models?

The questions and the related issues were answered in the study under different chapters utilising the research framework of literature review, data preparation, modelling and result analysis. The study, utilised different novel datasets both primary and secondary, for the city of Delhi, for example – survey data (Chapter 3), a spatial database of population, income, and caste for all residential locations in Delhi (Chapter 4), street map data (Chapter 5), land satellite imageries (Chapter 7). The study also utilised different geospatial models and statistical data techniques, for example, the geographically weighted regression model (Chapter 4), k-means clustering and SHAP method (Chapter 5), agent-based model (Chapter 6), and neural network coupled agent based model (Chapter 7).

The study findings showed that Delhi has a complex network of forms and functions which can be analysed at different spatial scales -individuals, households, neighbourhoods, and districts. Although the city core is ancient, its built-up form and transportation infrastructure is still evolving to cater for the need of incoming migrants. With increasing households' aspirations and income, the city requires to relook at the spatial planning of key services and needs to consider the neighbourhood population and socio-economic characteristics in the distribution of services across the city.

With the spatial expansion of metro services, the city has been able to provide higher mobility, however, the concern of accessibility still remains to be examined. The city will require a more comprehensive approach to cut down wasteful commuting and make the use of public transportation more attractive. While policies like transit-oriented development are important to enhance the use of public transportation, equally important is to bring a shift in the travel attitudes of commuters.

The street design remains one of the least examined elements of urban form in the city. It is important to revamp the existing streets to enhance walkability and neighbourhood vibrancy, especially in high-density neighbourhoods. Mixed land-use development in the neighbourhoods along with pedestrian plazas, and street greenery can help to enhance the active form of travel. Reducing traffic congestion in areas with high street intersection density remains a priority for the city administration, and it will require economic measures along with the use of information technology to manage the traffic flow.

Delhi is expected to urbanise in the coming years at a rapid pace. Managing the future built-up expansion in the city, planners need to prioritise the direction and nature of future growth, whether based on the likes of compact cities or sprawl development. The economic cost of housing and commuting does impact the residential location choice of households. Does controlling the residential plot price, housing rent price, and fares of public transportation, reshape the density pattern is something that needs to be examined for Delhi. If yes, economic measures can be used as a deterrent tool to put limits to the city's growth and manage urban expansion. Apart from the economic measures, planning measures related to transportation facilities and land use zoning can also be utilised to channel urban expansion. Recent advancement in geospatial modelling techniques is capable of predicting the spatial growth in built-up areas with good accuracy. Findings from such models shall be incorporated into the city master plans to build evidence-based urban policy.

8.5 Study Limitations and Future Scope of Work

(i) Study Limitations

We summarise here the important limitations of this study. First, the study to understand the impact of the built environment on commuting behaviour used the workplace as a destination but did not consider non-work-related commuting, which is also important to consider to understand the 'flow' component of the city. Second, while the household travel survey incorporated some subjective aspects of commuting behaviour, it did not very explicitly capture their daily travel activity routine. A more detailed questionnaire on the households' daily travel activity can provide more accurate insights into the factors affecting their commuting behaviour.

Third, the study while measuring the inequity in accessibility to services considers only the physical barriers. It is today widely recognised that individuals or a particular group even after having physical proximity, may not be able to access a service on grounds of economic unaffordability or social discrimination. A more realistic measure of inequity in accessibility should also consider the non-physical barriers to access to services. Fourth, the study while examining sustainable urbanisation considers only the impact of the neighbourhood's built-up form on sustainable living. Sustainability can also be examined from the environmental perspective with indicators such as air pollution, and vulnerability to extreme temperatures. With the increase in global warming, urban design can play an important role in lowering the surface temperature and thus, requires careful examination.

Finally, the study built an economic rational model to study the impact of city affordability on residential location choice. It thereby assumes that the individual decision to select a residential location is purely driven by economic reasons. However, it may not be so and other factors such as proximity to family or a particular neighbourhood may play an important role in determining their choice of residence and commuting. Thus, the impact of city affordability on residential location choice and commuting can be made more realistic by including the non-economic factors.

(ii) Future Scope of Work

While the study has explored some of the fundamental and unexamined issues in the city of Delhi, future work needs to be carried out to further understand the complexity of such systems. As cities are now understood to evolve through a bottom-up approach, the influence of individual actions and behaviour in shaping the urban form needs to be examined. With the advancement in machine learning and geo-computational techniques, future work on micro-simulation models on residential location choice and commuting behaviour will be of great significance in understanding how individual choices decide the density pattern and overall built-up form of the city.

Simulation based on re-enforcement learning or feedback loops can help us understand how the circular causality operates between individual choices and different urban form factors. For example, to understand the role of the built environment in influencing commuting behaviour, future studies need to consider the circular causality of the built environment, travel attitude, and commuting behaviour. As the built environment influences commuting behaviour, commuting behaviour in turn affects the built environment. While studies in transportation research till now have largely focused on the former, the latter remains unexplored. Such an understanding can help to design built environments that promote sustainable transportation and cut down wasteful commuting.

Simulation models can also be applied to understand the circular causality in the context of land price and accessibility to services. As cities grow, land with high accessibility to jobs and other services gets more in demand which enhances the land value and result of which, such land gets occupied by high-income households. Further, more services tend to be located near high-income neighbourhoods due to their high purchasing capacity and thus, such neighbourhoods tend to have high accessibility. In this manner, a vicious cycle sets up making the high-income neighbourhoods have higher accessibility to different services and segregating the low-income neighbourhoods. Using the mico-simulation models, the dynamic phenomena of residential segregation and inequity in accessibility can be studied and suitable policies to limit them can be framed.

Future studies in urban science also need to focus on incorporating planning measures in the simulation models and examining their impact on city growth. A top-down approach in urban planning, although still considered important, is not scientifically monitored, and evaluated. Many planned measures are designed to solve a specific issue without taking into account the possible negative spillover effects of such measures. Using simulation models one can study how a planned measure is going to change the system dynamics and under what conditions it can be executed to maximise its intended benefits.

As the world becomes increasingly urban, the sustainability paradigm in urban studies will gain more importance. Future studies need to suggest policy measures to make urbanisation, especially in the developing world more sustainable. That will require making the environment a stakeholder in designing and managing urban forms and functions.

References

Abascal, A., Rodríguez-Carreño, I., Vanhuysse, S., Georganos, S., Sliuzas, R., Wolff, E., & Kuffer, M. (2022). Identifying degrees of deprivation from space using deep learning and morphological spatial analysis of deprived urban areas. *Computers, environment and urban systems*, 95, 101820.

Abbott, A., Deshowitz, A., Murray, D., & Larson, E. C. (2018). WalkNet: A deep learning approach to improving sidewalk quality and accessibility. *SMU Data Science Review*, 1(1), 7.

Abdulai, I. A., Ahmed, A., & Kuusaana, E. D. (2022). Secondary cities under siege: examining peri-urbanisation and farmer households' livelihood diversification practices in Ghana. *Heliyon*, 8(9), e10540.

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11), e00938.

Abrantes, P., Rocha, J., Marques da Costa, E., Gomes, E., Morgado, P., & Costa, N. (2019). Modelling urban form: A multidimensional typology of urban occupation for spatial analysis. *Environment and Planning B: Urban Analytics and City Science*, 46(1), 47-65.

Abrantes, P., Rocha, J., Marques da Costa, E., Gomes, E., Morgado, P., & Costa, N. (2019). Modelling urban form: A multidimensional typology of urban occupation for spatial analysis. *Environment and Planning B: Urban Analytics and City Science*, 46(1), 47-65.

Acheampong, R. A. (2020). Spatial structure, intra-urban commuting patterns and travel mode choice: Analyses of relationships in the Kumasi Metropolis, Ghana. *Cities*, 96, 102432.

Acheampong, R. A., & Asabere, S. B. (2021). Simulating the co-emergence of urban spatial structure and commute patterns in an African metropolis: A geospatial agent-based model. *Habitat International*, 110, 102343.

Agyemang, F. S., & Silva, E. (2019). Simulating the urban growth of a predominantly informal Ghanaian city-region with a cellular automata model: Implications for urban planning and policy. *Applied Geography*, 105, 15-24.

Agyemang, F. S., Silva, E., & Fox, S. (2022). Modelling and simulating 'informal urbanization': An integrated agent-based and cellular automata model of urban residential growth in Ghana. *Environment and Planning B: Urban Analytics and City Science*, 23998083211068843.

Ahmad, S., & Choi, M. J. (2011). The context of uncontrolled urban settlements in Delhi. *Asien*, 118(January), 75-90.

Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, 9(8), 1295.

Ahmed, S., & Bramley, G. (2015). How will Dhaka grow spatially in future?-Modelling its urban growth with a near-future planning scenario perspective. *International Journal of Sustainable Built Environment*, 4(2), 359-377.

Ahrens, A., & Lyons, S. (2021). Do rising rents lead to longer commutes? A gravity model of commuting flows in Ireland. *Urban Studies*, 58(2), 264-279.

Ajzen, I. (1991). The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211.

Akaike, H. (1973). "Information Theory and an Extension of the Maximum Likelihood Principle." In BN Petrov, F Csaki (eds.), 2nd Symposium on Information Theory, pp. 267–281. Akademiai Kiado, Budapest

Alawadi, K., Khanal, A., Mouselly, A., & Aletaywi, A. B. (2022). Planning in the age of pandemics: Renewing suburban design. *Sustainable Cities and Society*, 87, 104261.

Alford, T. J., Shi, X., Onega, T., Wang, D., & Zou, J. (2008). Assess potential demand for major cancer care facilities in the US. In *Proceedings of the Association of American Geographers Annual Conference, Boston, MA*.

Alkamali, N., Alhadhrami, N., & Alalouch, C. (2017). Muscat City Expansion and Accessibility to the Historical Core: Space Syntax Analysis. *Energy Procedia*, 115: 480–486.

Allen, J., & Farber, S. (2019). Sizing up transport poverty: A national scale accounting of low-income households suffering from inaccessibility in Canada, and what to do about it. *Transport policy*, 74: 214-223.

Allen, J., & Farber, S. (2020). Planning transport for social inclusion: An accessibility-activity participation approach. *Transportation Research Part D: Transport and Environment*, 78, 102212.

Alonso, W. (1964). Location and land use: toward a general theory of land rent. *Harvard university press.*

AlQuhtani, S. & Anjomani, A. (2021). Do Rail Transit Stations Affect the Population Density Changes around Them? The Case of Dallas-Fort Worth Metropolitan Area. *Sustainability*, 13(6): 3355.

Al-Sayed, K., & Penn, A. (2017). Translating Analytical Descriptions of Cities into Planning and Simulation Models. *Design Computing and Cognition*, 16: 537–554.

Altieri, M., Silva, C., & Terabe, S. (2020). Give public transit a chance: A comparative analysis of competitive travel time in public transit modal share. *Journal of Transport Geography*, 87, 102817.

An, R., Wu, Z., Tong, Z., Qin, S., Zhu, Y., & Liu, Y. (2022). How the built environment promotes public transportation in Wuhan: A multiscale geographically weighted regression analysis. *Travel Behaviour and Society*, 29, 186-199.

Anderson, W. P., Kanargoglou, P. S. & Miller, E. (1996) Urban Form, Energy and theEnvironment: A Review of Issues, Evidence and Policy. *Urban Studies*, 33, pp. 17–35

Aram, F., García, E. H., Solgi, E., & Mansournia, S. (2019). Urban green space cooling effect in cities. *Heliyon*, 5(4), e01339.

Ashik, F. R., Mim, S. A., & Neema, M. N. (2020). Towards vertical spatial equity of urban facilities: An integration of spatial and aspatial accessibility. *Journal of Urban Management*, 9(1), 77-92.

Ashik, F. R., Rahman, M. H., Antipova, A., & Zafri, N. M. (2023). Analyzing the impact of the built environment on commuting-related carbon dioxide emissions. *International Journal of Sustainable Transportation*, 17(3), 258-272.

Babakan, A., & Alimohammadi, A. (2016). An agent-based simulation of residential location choice of tenants in Tehran, Iran. *Transactions in GIS*, 20(1), 101-125.

Babakan, A., & Taleai, M. (2015). Impacts of transport development on residence choice of renter households: An agent-based evaluation. *Habitat International*, 49, 275–285. doi:10.1016/j.habitatint.2015.05.033

Badhan, I. M. (2019). Space syntax analysis: Tracing the rationale for accessibility of recreational/movement economy growth along Hatirjheel Lake Park through integration and connectivity. *International Journal of Scientific & Engineering Research*, 10(10), 1197-1206.

Banister, D. (1994). Equity and acceptability in internalising the social costs of transport. *EMCT, Internalising the social costs of transport*, 153-175.

Baobeid, A., Koç, M., & Al-Ghamdi, S. G. (2021). Walkability and its relationships with health, sustainability, and livability: elements of physical environment and evaluation frameworks. *Frontiers in Built Environment*, 7, 721218.

Baraklianos, I., Bouzouina, L., Bonnel, P., & Aissaoui, H. (2020). Does the accessibility measure influence the results of residential location choice modelling? *Transportation*, 47(3), 1147-1176.

Basso, L. J., Navarro, M., & Silva, H. E. (2021). Public transport and urban structure. *Economics of Transportation*, 28, 100232.

Basu, R., & Ferreira, J. (2021). Planning car-lite neighborhoods: Does bikesharing reduce auto-dependence?. *Transportation research part D: transport and environment*, 92, 102721.

Batty, M. (2005). Agents, cells, and cities: new representational models for simulating multiscale urban dynamics. *Environment and Planning A*, 37(8), 1373-1394.

Batty, M. (2007). Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals. *The MIT Press.*

Batty, M. (2008). The size, scale, and shape of cities. *science*, 319(5864), 769-771.

Batty, M. (2009). Accessibility: in search of a unified theory. *Environment and Planning B: Planning and Design*, 36(2), 191-194.

Batty, M. (2013). The new science of cities. *MIT press.*

Becker, W. R., Ló, T. B., Johann, J. A., & Mercante, E. (2021). Statistical features for land use and land cover classification in Google Earth Engine. *Remote Sensing Applications: Society and Environment*, 21, 100459.

Ben-Akiva, M. E., & Lerman, S. R. (1985). Discrete choice analysis: theory and application to travel demand (Vol. 9). *MIT press.*

Bettencourt, L. M. (2021). Introduction to urban science: evidence and theory of cities as complex systems.

Bhatt, D. (2022). GIS and Gravity Model-Based Accessibility Measure for Delhi Metro. *Iranian Journal of Science and Technology*, Transactions of Civil Engineering, 46(4), 3411-3428.

Bibri, S. E. (2021). Data-driven smart sustainable cities of the future: An evidence synthesis approach to a comprehensive state-of-the-art literature review. *Sustainable Futures*, 3, 100047.

Bibri, S. E., Krogstie, J., & Kärrholm, M. (2020). Compact city planning and development: Emerging practices and strategies for achieving the goals of sustainability. *Developments in the built environment*, 4, 100021.

Bikis, A. (2023). Quantifying and analyzing the impact assessment on land use change of urban growth using a timeline. *Environmental Science and Pollution Research*, 1-20.

Bobkova, E., Berghauser Pont, M., & Marcus, L. (2021). Towards analytical typologies of plot systems: Quantitative profile of five European cities. *Environment and Planning B: Urban Analytics and City Science*, 48(4), 604-620.

Bourdic, L., Salat, S., & Nowacki, C. (2012). Assessing cities: a new system of cross-scale spatial indicators. *Building Research & Information*, 40(5), 592-605.

Bradley, M., Bowman, J. L., & Griesenbeck, B. (2010). SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution. *Journal of Choice Modelling*, 3(1), 5-31.

Brandsæter, A., & Glad, I. K. (2022). Shapley values for cluster importance. *Data Mining and Knowledge Discovery*, 1-32.

Braulio-Gonzalo, M., Ruá, M. J., & Bovea, M. D. (2018). Exploring residential urban form patterns: a Spanish case study. *International Planning Studies*.

Brown, D., Japa, A., & Shi, Y. (2019). A fast density-grid based clustering method. In *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0048-0054). IEEE.

Brueckner, J. K. (1987). The structure of urban equilibria: A unified treatment of the Muth-Mills model. *Handbook of regional and urban economics*, *2*(20), 821-845.

Brueckner, J.K., Thisse, J-F, & Zenou, Y. (1999). Why is central Paris rich and downtown detroit poor?: an amenity-based theory. *Eur Econ Rev*, 43(1), 91–107.

Brunsdon C., Fotheringham AS., & Charlton M. (1996). Geographically weighted regression: a method for exploring spatial non-stationarity. *Geogr Anal*, 28(4), 281–298.

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(3), 219-229.

Burgess, E. W. (1928). Residential segregation in American cities. *The annals of the american academy of political and social science*, 140(1), 105-115.

Burns L. D., Golob T. F. (1976). The role of accessibility in basic transportation choice behavior. *Transportation*, 5(2): 175–198.

Burns, L. D. (1980). Transportation, temporal, and spatial components of accessibility.

Cairncross, F. (1997). The death of distance: How the communications revolution will change our lives. (No Title).

Campbell, Jr. H. S. (2021). Income and cost of living: Are less equal places more costly? *Social Science Quarterly*, 102(6), 2689-2705.

Carr, D., Giddens, A., Duneier, M., & Appelbaum, R. P. (2018). Introduction to sociology (p. 575). *New York: WW Norton & Company*.

Carrella, E., Bailey, R., & Madsen, J. (2020). Calibrating agent-based models with linear regressions. Journal of Artificial Societies and Social Simulation, 23(1).

Caubel, D. (2005). Infra-municipal territorial disparities (IRIS-2000) according to living standards and social positions in the urban areas of Lyon, Bordeaux, Paris, Toulouse, Dijon, Pau, Agen and Villefranche-sur-Saône. In *XLIth ASRDLF Colloquium "Cities and territories facing the challenges of globalization*" (p. 20).

Chang, Z., Chen, J., Li, W., & Li, X. (2019). Public transportation and the spatial inequality of urban park accessibility: New evidence from Hong Kong. *Transportation Research Part D: Transport and Environment,* 76, 111-122.

Charles, M., & Lundy, J. D. (2013). The local Joneses: Household consumption and income inequality in large metropolitan areas. *Research in Social Stratification and Mobility*, 34, 14-29.

Chatterjee, K., Chng, S., Clark, B., Davis, A., De Vos, J., Ettema, D., ... & Reardon, L. (2020). Commuting and wellbeing: a critical overview of the literature with implications for policy and future research. *Transport reviews*, 40(1), 5-34.

Chaudhry, A. G., Masoumi, H., & Dienel, H. L. (2023). A systematic literature review of mobility attitudes and mode choices: MENA and South Asian cities. *Frontiers in Sustainable Cities*, 4, 1085784.

Chen, T. L., Chiu, H. W., & Lin, Y. F. (2020). How do east and southeast Asian cities differ from western cities? A systematic review of the urban form characteristics. *Sustainability*, 12(6), 2423.

Chen, Y. J., & Akar, G. (2017). Using trip chaining and joint travel as mediating variables to explore the relationships among travel behavior, socio-demographics, and urban form. *Journal of Transport and Land use*, 10(1), 573-588.

Chen, Y., Yue, W., & La Rosa, D. (2020). Which communities have better accessibility to green space? An investigation into environmental inequality using big data. *Landscape and Urban Planning*, 204, 103919.

Chen, Y.,et al. (2014). Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy. *International Journal of Geographical Information Science*, 28 (2), 234–255. doi:10.1080/13658816.2013.831868

Cheng J., Yan R., & Gao Y. (2020). Exploring spatial heterogeneity in accessibility and transit mode choice. *Transportation Research Part D: Transport and Environment*, 87.

Cheng, G., Zeng, X., Duan, L., Lu, X., Sun, H., Jiang, T., & Li, Y. (2016). Spatial difference analysis for accessibility to high level hospitals based on travel time in Shenzhen, China. *Habitat International*, 53: 485–494

Cheng, L., De Vos, J., Shi, K., Yang, M., Chen, X., & Witlox, F. (2019). Do residential location effects on travel behavior differ between the elderly and younger adults?. *Transportation research part D: transport and environment*, 73, 367-380.

Cheng, S., Fan, W., Zhang, J., Wang, N., Meng, F., & Liu, G. (2021). Multi-sectoral determinants of carbon emission inequality in Chinese clustering cities. *Energy*, 214, 118944.

Chidambaram, B., & Scheiner, J. (2021). Work-trip mode choice in Germany–Affected by individual constraints or by partner interaction?. *Travel Behaviour and Society*, 24, 231-244.

Choi, K., Lim, W., Chang, B., Jeong, J., Kim, I., Park, C. R., & Ko, D. W. (2022). An automatic approach for tree species detection and profile estimation of urban street trees using deep learning and Google street view images. *Isprs Journal of Photogrammetry and Remote Sensing*, 190, 165-180.

Clark Labs. (2018). About Clark Labs. *Clark Labs*. https://clarklabs.org/. Accessed 5 Mar 2018

Collier, P., Jones, P., & Spijkerman, D. (2018). Cities as engines of growth: Evidence from a new global sample of cities. Unpublished.

Consumer Expenditure Surveys, U.S. Bureau of Labor Statistics, September, 2021

Corrigan, E., Foley, D., McQuinn, K., O’Toole, C., & Slaymaker, R. (2019). Exploring affordability in the Irish housing market. *The Economic and Social Review*, 50(1), 119-157.

Cortés, Y. (2021). Spatial accessibility to local public services in an unequal place: An Analysis from patterns of residential segregation in the Metropolitan Area of Santiago, Chile. *Sustainability*, 13(2), 442.

Crooks, A.T. & Wise, S. (2013). GIS and Agent-Based models for Humanitarian Assistance. *Computers, Environment and Urban Systems*, 41, 100–111. [doi:10.1016/j.compenvurbsys.2013.05.003]

Crooks, AT. (2010) Constructing and implementing an agent-based model of residential segregation through vector GIS. *Int J Geogr Inf Sci*, 24(5), 661–675

Currie, G., & Senbergs, Z. (2007). Exploring forced car ownership in metropolitan Melbourne. In *Australasian transport research forum* (atrf), Melbourne, Victoria, Australia, vol 30.

Dadashpoor, H., Azizi, P., & Moghadasi, M. (2019). Analyzing spatial patterns, driving forces and predicting future growth scenarios for supporting sustainable urban growth: Evidence from Tabriz metropolitan area, Iran. *Sustainable Cities and Society*, 47, 101502.

Dahlke, J., & Bogner, M. et al. (2020). Is the Juice Worth the Squeeze? Machine Learning (ML) In and For Agent-Based Modelling (ABM). arXiv preprint arXiv:2003.11985.

Dalvi, M. Q., & Martin, K. M. (1976). The measurement of accessibility: Some preliminary results. *Transportation*, 5(1): 17–42.

Dano, U. L., Balogun, A. L., Abubakar, I. R., & Aina, Y. A. (2020). Transformative urban governance: Confronting urbanization challenges with geospatial technologies in Lagos, Nigeria. *GeoJournal*, 85, 1039-1056.

Davidoff, P. (1965). Advocacy and pluralism in planning. Journal of the American Institute of planners, 31(4), 331-338.

De Vos, J., & Witlox, F. (2013). Transportation policy as spatial planning tool; reducing urban sprawl by increasing travel costs and clustering infrastructure and public transportation. Journal of *Transport Geography*, 33, 117-125.

Debrah, C., Owusu-Manu, D. G., Amonoo-Parker, L., Baiden, B. K., Oduro-Ofori, E., & Edwards, D. J. (2022). A factor analysis of the key sustainability content underpinning green cities development in Ghana. International Journal of Construction Management, 1-10.

Debray, H., Kraff, N. J., Zhu, X. X., & Taubenböck, H. (2023). Planned, unplanned, or in-between? A concept of the intensity of plannedness and its empirical relation to the built urban landscape across the globe. *Landscape and Urban Planning*, 233, 104711.

Deka, D. (2015). Relationship between Households' Housing and Transportation Expenditures: Examination from Lifestyle Perspective. *Transportation Research Record*, 2531(1), 26-35.

Delamater, P. L. (2013). Spatial accessibility in suboptimally configured health care systems: A modified two-step floating catchment area (M2SFCA) metric. *Health & place*, 24, 30-43.

Delbosc, A., & Currie, G. (2018). Accessibility and exclusion related to well being. *Quality of life and daily travel*, 57-69.

Delhi Development Authority. (2022). List of Unauthorised colonies in Delhi. Accessed from url: https://dda.org.in/planningmap.aspx

Delmelle, E., Nilsson, I., & Adu, P. (2021). Poverty suburbanization, job accessibility, and employment outcomes. *Social Inclusion*, 9(2), 166.

Dempsey, N., Bramley, G., Power, S., & Brown, C. (2011). The social dimension of sustainable development: Defining urban social sustainability. *Sustainable development*, 19(5), 289-300.

Dempsey, N., Brown, C., & Bramley, G. (2012). The key to sustainable urban development in UK cities? The influence of density on social sustainability. *Progress in planning*, 77(3), 89-141.

Deng, Y., & Srinivasan, S. (2016). Urban land use change and regional access: A case study in Beijing, China. *Habitat international*, 51, 103-113.

Deng, Y., & Zhao, P. (2022). Quantifying residential self-selection effects on commuting mode choice: A natural experiment. *Transportation Research Part D: Transport and Environment*, 104, 103197.

Dewita, Y., Burke, M., & Yen, B. T. (2020). The relationship between transport, housing and urban form: Affordability of transport and housing in Indonesia. *Case Studies on Transport Policy*, 8(1), 252-262.

Dikeç, M. (2009). Space, politics and (in) justice. Justice spatiale-spatial justice

Ding, C., & Cao, X. (2019). How does the built environment at residential and work locations affect car ownership? An application of cross-classified multilevel model. *Journal of Transport Geography*, 75, 37-45.

Ding, C., Liu, T., Cao, X., & Tian, L. (2022). Illustrating nonlinear effects of built environment attributes on housing renters' transit commuting. *Transportation Research Part D: Transport and Environment*, 112, 103503.

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, T., Jiao, L., Xu, G., Yang, L., & Liu, J. (2019). Towards sustainability? Analyzing changing urban form patterns in the United States, Europe, and China. *Science of the Total Environment*, 671, 632-643.

Dulal, H. B. (2017). Making cities resilient to climate change: Identifying "win–win" interventions. *Local Environment*, 22(1), 106-125.

Duquet, B., & Brunelle, C. (2020). Subcentres as Destinations: Job Decentralization, Polycentricity, and the Sustainability of Commuting Patterns in Canadian Metropolitan Areas, 1996–2016. *Sustainability*, 12(23), 9966.

Duranton, G., & Puga, D. (2020). The economics of urban density. *Journal of economic perspectives*, 34(3), 3-26.

Dutta, D., Rahman, A., Paul, S. K., & Kundu, A. (2020). Estimating urban growth in peri-urban areas and its interrelationships with built-up density using earth observation datasets. *The Annals of Regional Science*, 65, 67-82.

Eastman, J. R. (2016). TerrSet geospatial monitoring and modeling system. *Clark University: Worcester, MA, USA*, 345-389.

Eckerd, A., Kim, Y., & Campbell, H. (2018). Gentrification and Displacement: Modeling a Complex Urban Process. *Housing Policy Debate*, 1–23.

Egidi, G., Cividino, S., Vinci, S., Sateriano, A., & Salvia, R. (2020). Towards local forms of sprawl: A brief reflection on mediterranean urbanization. *Sustainability*, 12(2), 582.

Ekanayake, I. U., Meddage, D. P. P., & Rathnayake, U. (2022). A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). *Case Studies in Construction Materials*, 16, e01059.

Engelfriet, L., & Koomen, E. (2018). The impact of urban form on commuting in large Chinese cities. *Transportation*, 45(5), 1269-1295.

Eurostat. (2018). Methodological manual on territorial typologies.

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment: A Meta-Analysis. *Journal of the American Planning Association*, 76(3), pp.265-294.

Fahim, A. (2021). K and starting means for k-means algorithm. *Journal of Computational Science*, 55, 101445.

Farahani, L. M., Izadpanahi, P., & Tucker, R. (2022). The death and life of Australian suburbs: Relationships between social activity and the physical qualities of Australian suburban neighbourhood centres. *City, Culture and Society*, 28, 100426.

Feilong, H., Shijun, W., Dongcan, X., Tingting, Y., & Zhangxian, F. (2017). Space-Time Accessibility of Commercial Centers in Changchun Urban Area Based on Internet Map Service. Econ. *Geogr*, 37: 68–75.

Feng X., Gan T., & Wang X. (2015). Feedback Analysis of Interaction between Urban Densities and Travel Mode Split. *International Journal of Simulation Modelling*, 14(2): 349–358.

Fleischmann, M., Feliciotti, A., Romice, O., & Porta, S. (2022). Methodological foundation of a numerical taxonomy of urban form. *Environment and Planning B: Urban Analytics and City Science*, 49(4), 1283-1299.

Fleischmann, M., Romice, O., & Porta, S. (2021). Measuring urban form: Overcoming terminological inconsistencies for a quantitative and comprehensive morphologic analysis of cities. *Environment and Planning B: Urban Analytics and City Science*, 48(8), 2133-2150.

Fosset, P., Banos, A., Beck, E., Chardonnel, S., Lang, C., & Marilleau N., Thevenin T. (2016). Exploring Intra-Urban Accessibility and Impacts of Pollution Policies with an Agent-Based Simulation Platform: GaMiroD. *Systems*, 4(1): 5.

Fotheringham, A. S. (1986). Modelling Hierarchical Destination Choice. *Environment and Planning A*, 18(3): 401–418.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). Geographically weighted regression: the analysis of spatially varying relationships. *John Wiley & Sons*.

Fotheringham, A. S., Nakaya, T., Yano, K., Openshaw, S., & Ishikawa, Y. (2001). Hierarchical Destination Choice and Spatial Interaction Modelling: A Simulation Experiment. *Environment and Planning A*, 33(5): 901–920.

François, J. C, Ribardière, A., Fleury, A., Mathian, H., Pavard, A., & Saint-Julien, T. (2011). Income disparities among Ile-de-France households – Analysis of changes between 1999 and 2007.

Fransen, K., Neutens, T., 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.

Fränti, P., & Sieranoja, S. (2019). How much can k-means be improved by using better initialization and repeats?. *Pattern Recognition*, 93, 95-112.

Fu, X., Lam, W. H. K., Chen, B. Y., & Liu, Z. (2020) Maximizing space-time accessibility in multi-modal transit networks: an activity-based approach. *Transportmetrica A: Transport Science*, 1–32.

Fu, X., Zuo, Y., Zhang, S., & Liu, Z. (2022). Measuring joint space-time accessibility in transit network under travel time uncertainty. *Transport Policy*, 116, 355-368.

Fusco, G., Araldi, A., Emsellem, D., & Overal, D. (2022). Taxonomy of contemporary urban forms in France: Towards an Urban Atlas through Multiple Fabric Assessment. In *XXIX International Seminar on Urban Form ISUF2022: Urban Redevelopment and Revitalisation*. A Multidisciplinary Perspective-Proceedings.

Gabriel, M., Jacobs, K., Arthurson, K., Burke, T., & Yates, J. (2005). Conceptualising and Measuring the Housing Affordability Problem, National Research Venture 3: Housing Affordability for Lower Income Australians. *Research paper, 1.*

Gaglione, F., Cottrill, C., & Gargiulo, C. (2021). Urban services, pedestrian networks and behaviors to measure elderly accessibility. *Transportation research part D: transport and environment*, 90, 102687.

Galaskiewicz, J., Anderson, K. F., & Thompson-Dyck, K. (2021). Minority-White income inequality across metropolitan areas: The role of racial/ethnic residential segregation and transportation networks. *Journal of Urban Affairs*, 43(1), 16-39.

García-Albertos, P., Picornell, M., Salas-Olmedo, M. H., & Gutiérrez, J. (2019). Exploring the potential of mobile phone records and online route planners for dynamic accessibility analysis. *Transportation Research Part A: Policy and Practice*, 125, 294-307.

Gaur, S., & Singh, R. (2023). A Comprehensive Review on Land Use/Land Cover (LULC) Change Modeling for Urban Development: Current Status and Future Prospects. *Sustainability*, 15(2), 903.

GEE. (2021). Google Earth Engine Landsat Algorithms [WWW Docu-ment]. Simple cloud score algorithm. https://developers.google.com/earth-engine/landsat#simplecloud-score.

Gelormino, E., Melis, G., Marietta, C., & Costa, G. (2015). From built environment to health inequalities: An explanatory framework based on evidence. *Preventive medicine reports*, 2, 737-745.

Geurs, K. T., & Van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, 12(2): 127-140.

Gharaibeh, A., Shaamala, A., Obeidat, R., & Al-Kofahi, S. (2020). Improving land-use change modeling by integrating ANN with Cellular Automata-Markov Chain model. Heliyon, 6(9), e05092.

Gillespie, A. E. (1992). Communications technologies and the future of the city. In 22nd Annual Conference of the British Section of the Regional Science Association. Newcastle University.

Girma, R., Fürst, C., & Moges, A. (2022). Land use land cover change modeling by integrating artificial neural network with cellular Automata-Markov chain model in Gidabo river basin, main Ethiopian rift. *Environmental challenges*, 6, 100419.

Glaeser, E.L., Kahn, M.E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics* 63, 1– 24.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, 63(17).

Gore, R. J., Lynch, C. J., & Kavak, H. (2017). Applying statistical debugging for enhanced trace validation of agent-based models. *SIMULATION*, 93(4): 273–284.

Gottmann, J. (1983). Urban settlements and telecommunications. Ekistics, 411-416.

Govender, P., & Sivakumar, V. (2020). Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric pollution research*, 11(1), 40-56.

Government of Delhi. (2014). Notification on circle rate. Revenue Department, National Captial Territory of Delhi. Accessed via: https://srams.delhi.gov.in/docs_avail/c_28ab375a72f800406822233e3497e649_ircle1.pdf

Goytia, C., Heikkila, E. J., & Pasquini, R. A. (2023). Do land use regulations help give rise to informal settlements? Evidence from Buenos Aires. *Land Use Policy*, 125, 106484.

Guagliardo, M. F. (2004). Spatial accessibility of primary care: concepts, methods and challenges. *International journal of health geographics*, 3(1), 1-13.

Guan C, & Peiser R. (2018). Accessibility, urban form, and property value: A study of Pudong, Shanghai. *Journal of Transport and Land Use*, 11(1).

Guan, C., Srinivasan, S., & Nielsen, C. P. (2019). Does neighborhood form influence low-carbon transportation in China? *Transportation Research Part D: Transport and Environment*, 67: 406–420.

Guan, X., & Wang, D. (2019). Residential self-selection in the built environment-travel behavior connection: Whose self-selection?. *Transportation Research Part D: Transport and Environment*, 67, 16-32.

Guan, X., Wang, D., & Jason Cao, X. (2020). The role of residential self-selection in land use-travel research: a review of recent findings. *Transport reviews*, 40(3), 267-287.

Guerra, E., Li, S., & Reyes, A. (2022). How do low-income commuters get to work in US and Mexican cities? A comparative empirical assessment. *Urban Studies*, 59(1), 75-96.

Guida, C., & Carpentieri, G. (2021). Quality of life in the urban environment and primary health services for the elderly during the Covid-19 pandemic: An application to the city of Milan (Italy). *Cities*, 110, 103038.

Guidon, S., Wicki, M., Bernauer, T., & Axhausen, K. (2019). The social aspect of residential location choice: on the trade-off between proximity to social contacts and commuting. *Journal of Transport Geography*, 74, 333-340.

Gul, Y., Jokhio, G. A., & Bibi, T. (2020). Walk towards sustainability: Improved neighbourhood street connectivity helps. In *E3S Web of Conferences* (Vol. 211, p. 01004). EDP Sciences.

Guo C., Buchmann C. M., Schwarz N. (2017). Linking urban sprawl and income segregation – Findings from a stylized agent-based model. *Environment and Planning B: Urban Analytics and City Science.*

Guo, S., Song, C., Pei, T., Liu, Y., Ma, T., Du, Y., … & Wang, Y. (2019). Accessibility to urban parks for elderly residents: Perspectives from mobile phone data. *Landscape and urban planning*, 191, 103642.

Gupta, A., Bivina, G. R., & Parida, M. (2022). Does neighborhood design matter for walk access to metro stations? An integrated SEM-Hybrid discrete mode choice approach. *Transport policy*, 121, 61-77.

Guyot, M., Araldi, A., Fusco, G., & Thomas, I. (2021). The urban form of Brussels from the street perspective: The role of vegetation in the definition of the urban fabric. *Landscape and Urban Planning*, 205, 103947.

Ha, T. V., Asada, T., & Arimura, M. (2019). Determination of the influence factors on household vehicle ownership patterns in Phnom Penh using statistical and machine learning methods. *Journal of Transport Geography*, 78, 70–86.

Haas, P. M., Makarewicz, C., Benedict, A., Sanchez, T. W., & Dawkins, C. J. (2006). Housing & transportation cost trade-offs and burdens of working households in 28 metros. *Center for Neighborhood Technology*, *2*.

Hägerstrand, T. (1975). Space, time and human conditions. *Dynamic allocation of urban space*, 3: 2-12.

Halás, M., Klapka, P., & Kladivo, P. (2014). Distance-decay functions for daily travel-to-work flows. *Journal of Transport Geography*, 35: 107–119.

Hamilton, B. W., & Röell, A. (1982). Wasteful commuting. *Journal of political economy*, 90(5), 1035-1053.

Han, Y., Zhong, T., Yeh, A. G., Zhong, X., Chen, M., & Lü, G. (2023). Mapping seasonal changes of street greenery using multi-temporal street-view images. *Sustainable Cities and Society*, 92, 104498.

Handy, S. (2020). Is accessibility an idea whose time has finally come?. *Transportation research part D: transport and environment*, 83.

Handy, S. L, & Niemeier, D. A. (1997). Measuring Accessibility: An Exploration of Issues and Alternatives. *Environment and Planning A*, 29(7): 1175–1194.

Hankey S., & Marshall J. D. (2017). Urban form, air pollution, and health Curr. Environ. *Health Rep*. 4, 491–503

Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, 25(2): 73–76.

Harvey, D. (1973). Social Justice and the City. Oxford: Blackwell.

Hashemi, Aslani, Z., Omidvar, B., & Karbassi, A. (2022). Integrated model for land-use transformation analysis based on multi-layer perception neural network and agent-based model. *Environmental Science and Pollution Research*, 29(39), 59770-59783.

Havet, N., Bayart, C., & Bonnel, P. (2021). Why do Gender Differences in Daily Mobility Behaviours persist among workers? *Transp. Res. A Policy Pract*. 145, 34–48. https:// doi.org/10.1016/j.tra.2020.12.016.

Heikkila, E. J., & Harten, J. G. (2023). Can Land Use Regulation Be Smarter? Planners' Role in the Informal Housing Challenge. *Journal of Planning Education and Research*, 43(2), 359-370.

Henderson, J. V. (1974). The sizes and types of cities. *The American Economic Review*, 64(4), 640-656.

Heppenstall A., Crooks A., Malleson N., Manley E., Ge J., & Batty M. (2021). Future Developments in Geographical Agent-Based Models: Challenges and Opportunities. *Geogr. Anal*., 53: 76-91.

Heppenstall, A. & Malleson, N. (2020). Building Cities from Slime Mould, Agents and Quantum Field Theory. *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, edited by B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, and G. Sukthankar. Auckland, New Zealand.

Hillier, B., & Hanson, J. (1984). The Social Logic of Space. *Cambridge University Press*, Cambridge.

Hou, H., Estoque, R. C., & Murayama, Y. (2016). Spatiotemporal analysis of urban growth in three African capital cities: A grid-cell-based analysis using remote sensing data. *Journal of African Earth Sciences*, 123, 381-391.

Hu, L., He, S., Han, Z., Xiao, H., Su, S., Weng, M., & Cai, Z. (2019). Monitoring housing rental prices based on social media: An integrated approach of machine-learning algorithms and hedonic modeling to inform equitable housing policies. *Land use policy*, 82, 657-673.

Hu, Q., Ma, Y., Xu, B., Song, Q., Tang, H., & Wu, W. (2018). Estimating sub-pixel soybean fraction from time-series MODIS data using an optimized geographically weighted regression model. *Remote Sensing*, 10(4), 491.

Huang, A., & Xu, Y. (2005). Beijing urban redevelopment in the old center: do poors still can live there? *China news week*.

Huang, B. X., Chiou, S. C., & Li, W. Y. (2020). Accessibility and street network characteristics of urban public facility spaces: Equity research on parks in Fuzhou city based on GIS and space syntax model. *Sustainability*, 12(9), 3618.

Huang, R. (2017). Simulating individual work trips for transit-facilitated accessibility study. *Environment and Planning B: Urban Analytics and City Science*, 46(1): 84-102

Huang, X., & Chen, M. (2022). Understanding the role of housing in rural migrants' intention to settle in cities: Evidence from China. *Habitat International*, 128, 102650.

Hunter, E., Mac Namee, B., & Kelleher, J. D. (2018). A comparison of agent-based models and equation based models for infectious disease epidemiology.

Hurvich, C. M., Simonoff, J. S., & Tsai, C. L. (1998). "Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion." *Journal of the Royal Statistical Society B*, 60(2), 271–293.

Ingram, D. R. (1971). The concept of accessibility: A search for an operational form. *Regional Studies*, 5(2), 101–107.

Inturri, G., Ignaccolo, M., Le Pira, M., Caprì, S., & Giuffrida, N. (2017). Influence of accessibility, land use and transport policies on the transport energy dependence of a city. *Transportation research procedia*, *25*, 3273-3285.

Iqbal, N., Ravan, M., Jamshed, A., Birkmann, J., Somarakis, G., Mitraka, Z., & Chrysoulakis, N. (2022). Linkages between Typologies of Existing Urban Development Patterns and Human Vulnerability to Heat Stress in Lahore. *Sustainability*, 14(17), 10561.

Islam, K., Rahman, M. F., & Jashimuddin, M. (2018). Modeling land use change using cellular automata and artificial neural network: The case of Chunati Wildlife Sanctuary, Bangladesh. *Ecological indicators*, *88*, 439-453.

Islam, M. R., & Saphores, J. D. M. (2022). An LA story: The impact of housing costs on commuting. *Journal of Transport Geography*, 98, 103266.

Jacobs, J. (1961). Jane jacobs. The Death and Life of Great American Cities, 21(1), 13-25.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM computing surveys* (CSUR), 31(3), 264-323.

Jamtsho, S., Corner, R., & Dewan, A. (2015). Spatio-temporal analysis of spatial accessibility to primary health care in Bhutan. *ISPRS International Journal of Geo-Information*, *4*(3), 1584-1604.

Jiang, C., Li, J., & Liu, J. (2022). Does urbanization affect the gap between urban and rural areas? Evidence from China. *Socio-Economic Planning Sciences*, 82, 101271.

Jiawei, H., Micheline, K., & Jian, P. (2012). Data Mining: Concepts and Techniques.-3rd. Morgan kaufmann.

Jin, J. (2019). The effects of labor market spatial structure and the built environment on commuting behavior: Considering spatial effects and self-selection. *Cities*, *95*, 102392.

Jin, T., Cheng, L., Wang, K., Cao, J., Huang, H., & Witlox, F. (2022). Examining equity in accessibility to multi-tier healthcare services across different income households using estimated travel time. *Transport Policy*, 121, 1-13.

Jochem, W. C., Leasure, D. R., Pannell, O., et al. (2020). Classifying settlement types from multi-scale spatial patterns of building footprints. *Environment and Planning B: Urban Analytics and City Science*, 48, 1161–1179.

Johansson-Stenman, O., & Martinsson, P. (2006). Honestly, why are you driving a BMW?. *Journal of Economic Behavior & Organization*, *60*(2), 129-146.

Johnson, R. A., & Wichern, D. W. (2007). Applied multivariate statistical analysis (6th ed.). Prentice Hall.

Kahn, M. E. (2000). The environmental impact of suburbanization. *Journal of policy analysis and management*, 19(4), 569-586.

Kain, J. (1968) Housing segregation, Negro employment, and metropolitan decentralisation, The *Quarterly Journal of Economics*, 82, pp.175-197.

Kale, R. K., & Acharya, S. S. (2022). Introduction—Marginalization in India—Matters of Inclusion and Access. In *Mapping Identity-Induced Marginalisation in India: Inclusion and Access in the Land of Unequal Opportunities* (pp. 1-15). Singapore: Springer Nature Singapore.

Kamusoko, C., & Gamba, J. (2015). Simulating urban growth using a random forest-cellular automata (RF-CA) model. *ISPRS International Journal of Geo-Information*, *4*(2), 447-470.

Karji, A., Woldesenbet, A., Khanzadi, M., & Tafazzoli, M. (2019). Assessment of social sustainability indicators in mass housing construction: a case study of Mehr housing project. *Sustainable Cities and Society*, 50, 101697.

Kärmeniemi, M., Lankila, T., Ikäheimo, T., Koivumaa-Honkanen, H., & Korpelainen, R. (2018). The built environment as a determinant of physical activity: a systematic review of longitudinal studies and natural experiments. *Annals of behavioral medicine*, 52(3), 239-251.

Kasraian, D., Maat, K., & van Wee, B. (2019). The impact of urban proximity, transport accessibility and policy on urban growth: A longitudinal analysis over five decades. *Environment and Planning B: Urban Analytics and City Science*, *46*(6), 1000-1017.

Kasraian, D., Raghav, S., & Miller, E. J. (2020). A multi-decade longitudinal analysis of transportation and land use co-evolution in the Greater Toronto-Hamilton Area. *Journal of Transport Geography*, *84*, 102696.

Ke, X., Zheng, W., Zhou, T., & Liu, X. (2017). A CA-based land system change model: LANDSCAPE. *International Journal of Geographical Information Science*, *31*(9), 1798-1817.

Kellett, J., Morrissey, J., & Karuppannan, S. (2012). The impact of location on housing affordability. In *6th Australasian housing researchers' conference. Adelaide: The University of Adelaide.*

Kelobonye, K., McCarney, G., Xia, J. C., Swapan, M. S. H., Mao, F., & Zhou, H. (2019). Relative accessibility analysis for key land uses: A spatial equity perspective. *Journal of Transport Geography*, *75*, 82-93.

Kersting, M., Matthies, E., Lahner, J., & Schlüter, J. (2021). A socioeconomic analysis of commuting professionals. *Transportation*, 48, 2127-2158.

Khadem, H., Nemat, H., Eissa, M. R., Elliott, J., & Benaissa, M. (2022). COVID-19 mortality risk assessments for individuals with and without diabetes mellitus: Machine learning models integrated with interpretation framework. *Computers in Biology and Medicine*, 144, 105361.

Khan, A. A. (1992). An integrated approach to measuring potential spatial access to health care services. *Socio-economic planning sciences*, *26*(4), 275-287.

Khan, A. A., & Bhardwaj, S. M. (1994). Access to health care: a conceptual framework and its relevance to health care planning. *Evaluation & the health professions*, *17*(1), 60-76.

Khavarian-Garmsir, A. R., Sharifi, A., & Sadeghi, A. (2023). The 15-minute city: Urban planning and design efforts toward creating sustainable neighborhoods. *Cities*, 132, 104101.

Kim, H. M., & Kwan, M. P. (2003). Space-time accessibility measures: A geocomputational algorithm with a focus on the feasible opportunity set and possible activity duration. *Journal of geographical Systems*, *5*, 71-91.

Kim, Y., Newman, G., & Güneralp, B. (2020). A review of driving factors, scenarios, and topics in urban land change models. *Land*, 9(8), 246.

Kinigadner, J., Büttner, B., Wulfhorst, G., & Vale, D. (2020). Planning for low carbon mobility: Impacts of transport interventions and location on carbon-based accessibility. *Journal of Transport Geography*, *87*, 102797.

Kinigadner, J., Vale, D., Büttner, B., & Wulfhorst, G. (2021). Shifting perspectives. *Journal of Transport and Land Use*, *14*(1), 345-365.

Knowles, R. D., & Ferbrache, F. (Eds.). (2019). *Transit Oriented Development and sustainable cities*. Edward Elgar Publishing.

Ko, J., Lee, S., & Byun, M. (2019). Exploring factors associated with commute mode choice: An application of city-level general social survey data. *Transport policy*, 75, 36-46.

Kofinas, P., Dounis, A. I., & Vouros, G. A. (2018). Fuzzy Q-Learning for multi-agent decentralized energy management in microgrids. *Applied energy*, *219*, 53-67.

Koopmans, C., Rietveld, P., & Huijg, A. (2012). An accessibility approach to railways and municipal population growth, 1840–1930. *Journal of Transport Geography*, *25*, 98-104.

Kraff, N. J., Wurm, M., & Taubenböck, H. (2020). The dynamics of poor urban areas-analyzing morphologic transformations across the globe using Earth observation data. *Cities*, 107, 102905.

Kroesen, M., & Chorus, C. (2018). The role of general and specific attitudes in predicting travel behavior–A fatal dilemma?. *Travel behaviour and society*, 10, 33-41.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, 99(3), 483-499.

Kuddus, M. A., Tynan, E., & McBryde, E. (2020). Urbanization: a problem for the rich and the poor?. *Public health reviews*, 41, 1-4.

Kulish, M., Richards, A., Gillitzer, C. (2012). Urban structure and housing prices: some evidence from Australian cities. *Econ Rec*, 88(282), 303–322

Kundu, S., Khare, D., & Mondal, A. (2017) Past, present and future land-use changes and their impact on water balance. *J Environ Manage*, 197:582–596

Kushwaha, S., & Nithiyanandam, Y. (2019). The Study of Heat Island and its Relation With Urbanization in Gurugram, DELHI NCR for the Period of 1990 to 2018. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 49-56.

Kusuma, Y. S., Kaushal, S., Sundari, A. B., & Babu, B. V. (2018). Access to childhood immunisation services and its determinants among recent and settled migrants in Delhi, India. *Public Health*, 158, 135-143.

Kwan, M. P., & Kotsev, A. (2015). Gender differences in commute time and accessibility in S ofia, B ulgaria: a study using 3 D geovisualisation. *The Geographical Journal*, *181*(1), 83-96.

Laite, R., Portman, N., & Sankaranarayanan, K. (2016, December). Behavioral analysis of agent based service channel design using neural networks. In *2016 Winter Simulation Conference (WSC)* (pp. 3694-3695). IEEE.

Lamperti, F., Roventini, A., & Sani, A. (2018). Agent-based model calibration using machine learning surrogates. *Journal of Economic Dynamics and Control*, *90*, 366-389.

Langford, M., Higgs, G., & Fry, R. (2016). Multi-modal two-step floating catchment area analysis of primary health care accessibility. *Health & place*, *38*, 70-81.

Lazri, M., & Ameur, S. (2018). Combination of support vector machine, artificial neural network and random forest for improving the classification of convective and stratiform rain using spectral features of SEVIRI data. *Atmospheric research*, 203, 118-129.

Leao, S. Z., Huynh, N., Taylor, A., Pettit, C., & Perez, P. (2017). Evolution of a synthetic population and its daily mobility patterns under spatial strategies for urban growth. *Planning Support Science for Smarter Urban Futures 15*, 399-417.

Leasure, D. R., Jochem, W. C., Weber, E. M., Seaman, V., & Tatem, A. J. (2020). National population mapping from sparse survey data: A hierarchical Bayesian modeling framework to account for uncertainty. *Proceedings of the National Academy of Sciences*, 117(39), 24173-24179.

Lecompte, M. C., & Pablo, B. S. J. (2017). Transport systems and their impact con gender equity. *Transportation research procedia*, *25*, 4245-4257.

Lee, D., Derrible, S., & Pereira, F. C. (2018). Comparison of four types of artificial neural network and a multinomial logit model for travel mode choice modeling. *Transportation Research Record*, *2672*(49), 101-112.

Lee, J. S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, F., Lorscheid, I., ... & Parker, D. C. (2015). The complexities of agent-based modeling output analysis. *Journal of Artificial Societies and Social Simulation*, *18*(4).

Lee, J., & Miller, H. J. (2019). Analyzing collective accessibility using average space-time prisms. *Transportation Research Part D: Transport and Environment*, *69*, 250-264.

Lee, J., He, S. Y., & Sohn, D. W. (2017). Potential of converting short car trips to active trips: The role of the built environment in tour-based travel. *Journal of Transport & Health*, *7*, 134-148.

Lee, S., Yoo, C., & Seo, K. W. (2020). Determinant factors of pedestrian volume in different land-use zones: Combining space syntax metrics with GIS-based built-environment measures. *Sustainability*, *12*(20), 8647.

Lefebvre, H. (1996 [1968]) "The right to the city," in H. Lefebvre, Writings on Cities, E. Kofinan and E. Lebas (eds), Cambridge, Mass.: Blackwell, pp. 63-184.

Lemoy, R., Raux, C., & Jensen, P. (2013). Where in cities do “rich” and “poor” people live? The urban economics model revisited. hal-00805116v3

Lennox, J. (2020). More working from home will change the shape and size of cities. Victoria University, Centre of Policy Studies/IMPACT Centre.

Lenntorp, B. (1976). Paths in time-space environments: A time-geographic study of movement possibilities of individuals: University of Lund.

Leung, Y., Mei, C. L., & Zhang, W. X. (2000). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A*, *32*(1), 9-32.

Li, F., Zhou, T., & Lan, F. (2021). Relationships between urban form and air quality at different spatial scales: A case study from northern China. *Ecological Indicators*, 121, 107029.

Li, H., Wei, Y. D., Wu, Y., & Tian, G. (2019). Analyzing housing prices in Shanghai with open data: Amenity, accessibility and urban structure. *Cities*, *91*, 165-179.

Li, J., Song, G., Semakula, H. M., Dou, Y., & Zhang, S. (2019). Food access inequalities in Chinese urban neighborhoods: a case study of the Dalian development zone. *Food Security*, 11, 1087-1099.

Li, X., & Gong, P. (2016). Urban growth models: progress and perspective. *Science bulletin*, *61*(21), 1637-1650.

Li, X., Liu, X., & Yu, L. (2014). A systematic sensitivity analysis of constrained cellular automata model for urban growth simulation based on different transition rules. *International Journal of Geographical Information Science*, *28*(7), 1317-1335.

Li, Y., Lin, Y., Geertman, S., Hooimeijer, P., & Xu, W. A. (2021). Accessibility-based equity of public facilities: A case study in Xiamen, China. *Applied Spatial Analysis and Policy*, 14(4), 947-968.

Li, Z., Han, Z., Xin, J., Luo, X., Su, S., & Weng, M. (2019). Transit oriented development among metro station areas in Shanghai, China: Variations, typology, optimization and implications for land use planning. *Land use policy*, 82, 269-282.

Li, Z., He, S., Su, S., Li, G., & Chen, F. (2020). Public services equalization in urbanizing China: Indicators, spatiotemporal dynamics and implications on regional economic disparities. *Social Indicators Research*, 152(1), 1-65.

Li, Z., Wang, X., Zarazaga, J., Smith-Colin, J., & Minsker, B. (2022). Do infrastructure deserts exist? Measuring and mapping infrastructure equity: A case study in Dallas, Texas, USA. *Cities*, 130, 103927.

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.

Liang, Y., Du, M., Wang, X., & Xu, X. (2020). Planning for urban life: A new approach of sustainable land use plan based on transit-oriented development. *Evaluation and program planning*, 80, 101811.

Lindert, P., & Westen, A. (1991). Household shelter strategies in comparative perspective: evidence from low-income groups in Bamako and La Paz. *World Development*, 19(8), 1007-1028.

Litman, T. (2002). Evaluating transportation equity. *World Transport Policy & Practice*, 8(2), 50-65.

Liu, C., Sun, Y., Chen, Y., & Susilo, Y. O. (2018). The effect of residential housing policy on car ownership and trip chaining behaviour in Hangzhou, China. *Transportation Research Part D: Transport and Environment*, 62, 125-138.

Liu, D., Kwan, M. P., & Kan, Z. (2021). Analysis of urban green space accessibility and distribution inequity in the City of Chicago. *Urban Forestry & Urban Greening*, 59, 127029.

Liu, D., Kwan, M. P., Kan, Z., & Song, Y. (2021). An integrated analysis of housing and transit affordability in the Chicago metropolitan area. *The Geographical Journal*, 187(2), 110-126.

Liu, J., Ettema, D., & Helbich, M. (2022). Systematic review of the association between commuting, subjective wellbeing and mental health. *Travel behaviour and society*, 28, 59-74.

Liu, X., Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., ... & Pei, F. (2017). A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning*, *168*, 94-116.

Liu, Y., Cao, X., & Li, T. (2020). Influence of accessibility on land use and landscape pattern based on mapping knowledge domains: review and implications. *Journal of Advanced Transportation*, *2020*, 1-12.

Liu, Z., Ma, J., & Chai, Y. (2017). Neighborhood-scale urban form, travel behavior, and CO2 emissions in Beijing: implications for low-carbon urban planning. *Urban Geography*, *38*(3), 381-400.

Lowe, J. S., Prochaska, N., & Keating, W. D. (2022). Bringing permanent affordable housing and community control to scale: The potential of Community Land Trust and land bank collaboration. *Cities*, 126, 103718.

Lowe, M., Adlakha, D., Sallis, J. F., Salvo, D., Cerin, E., Moudon, A. V., ... & Giles-Corti, B. (2022). City planning policies to support health and sustainability: an international comparison of policy indicators for 25 cities. *The Lancet global health*, 10(6), e882-e894.

Lu, J., Li, B., Li, H., & Al-Barakani, A. (2021). Expansion of city scale, traffic modes, traffic congestion, and air pollution. *Cities*, 108, 102974.

Lu, S., Huang, Y., Shi, C., & Yang, X. (2019). Exploring the associations between urban form and neighborhood vibrancy: a case study of Chengdu, China. *ISPRS International Journal of Geo-Information*, 8(4), 165.

Lu, S., Shi, C., & Yang, X. (2019). Impacts of built environment on urban vitality: Regression analyses of Beijing and Chengdu, China. *International journal of environmental research and public health*, *16*(23), 4592.

Lu, Y., Sun, G., Sarkar, C., Gou, Z., & Xiao, Y. (2018). Commuting mode choice in a high-density city: Do land-use density and diversity matter in Hong Kong?. *International journal of environmental research and public health*, *15*(5), 920.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.

Luo, J. (2014). Integrating the huff model and floating catchment area methods to analyze spatial access to healthcare services. *Transactions in GIS*, *18*(3), 436-448.

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*(4), 1100-1107.

Luo, W., & Whippo, T. (2012). Variable catchment sizes for the two-step floating catchment area (2SFCA) method. *Health & place*, *18*(4), 789-795.

Lynge, H., Visagie, J., Scheba, A., Turok, I., Everatt, D., & Abrahams, C. (2022). Developing neighbourhood typologies and understanding urban inequality: a data-driven approach. *Regional Studies, Regional Science*, 9(1), 618-640.

Määttä-Juntunen, H., Antikainen, H., Kotavaara, O., & Rusanen, J. (2011). Using GIS tools to estimate CO2 emissions related to the accessibility of large retail stores in the Oulu region, Finland. *Journal of transport geography*, *19*(2), 346-354.

Macal, C. M. (2016). Everything you need to know about agent-based modelling and simulation. *Journal of Simulation*, *10*, 144-156.

Mahmoudi, J., & Zhang, L. (2018). Impact of county-level built environment and regional accessibility on walking: A Washington, DC–Baltimore case study. *Journal of Urban Planning and Development*, *144*(3), 04018020.

Mahtta, R., Mahendra, A., & Seto, K. C. (2019). Building up or spreading out? Typologies of urban growth across 478 cities of 1 million+. *Environmental Research Letters*, 14(12), 124077.

Makarewicz, C., Dantzler, P., & Adkins, A. (2020). Another Look at Location Affordability: Understanding the Detailed Effects of Income and Urban Form on Housing and Transportation Expenditures. *Housing Policy Debate*, 30(6), 1033-1055.

Manson, S., An, L., Clarke, K. C., Heppenstall, A., Koch, J., Krzyzanowski, B., ... & Tesfatsion, L. (2020). Methodological issues of spatial agent-based models. *Journal of Artificial Societies and Social Simulation*, *23*(1).

Marini, M., Chokani, N., & Abhari, R. S. (2019). Immigration and future housing needs in Switzerland: Agent-based modelling of agglomeration Lausanne. *Computers, Environment and Urban Systems*, *78*, 101400.

Marshall, A. (1890). Principles of economics Macmillan. London

Marwal, A., & Silva, E. (2022). Literature review of accessibility measures and models used in land use and transportation planning in last 5 years. *Journal of Geographical Sciences*, 32(3), 560-584.

Masoumi, H. E., Terzi, F., & Serag, Y. M. (2019). Neighborhood-scale urban form typologies of large metropolitan areas: Observations on Istanbul, Cairo, and Tehran. *Cities*, 85, 170-186.

Mattingly, K., & Morrissey, J. (2014). Housing and transport expenditure: Socio-spatial indicators of affordability in Auckland. *Cities*, 38, 69-83.

Mayaud, J. R., Tran, M., & Nuttall, R. (2019). An urban data framework for assessing equity in cities: Comparing accessibility to healthcare facilities in Cascadia. *Computers, Environment and Urban Systems*, 78, 101401.

McFadden, D. (1978). Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments. In D.Hensher and P. Stopher (eds.), *Behavioural Travel Modelling*, 279-318, Croom Helm London: London.

McGrail, M. R., & Humphreys, J. S. (2014). Measuring spatial accessibility to primary health care services: Utilising dynamic catchment sizes. *Applied Geography*, *54*, 182-188.

Meddage, P., Ekanayake, I., Perera, U. S., Azamathulla, H. M., Md Said, M. A., & Rathnayake, U. (2022). Interpretation of machine-learning-based (black-box) wind pressure predictions for low-rise gable-roofed buildings using Shapley additive explanations (SHAP). *Buildings*, 12(6), 734.

Meerow, S., Pajouhesh, P., & Miller, T. R. (2019). Social equity in urban resilience planning. *Local Environment*, 24(9), 793-808.

Mehrotra, S., Bardhan, R., & Ramamritham, K. (2019). Outdoor thermal performance of heterogeneous urban environment: An indicator-based approach for climate-sensitive planning. *Science of the Total Environment*, 669, 872-886.

Miller, H. J. (1991). Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information System*, *5*(3), 287-301.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, 57(2), 197-210.

Mishra, D., & Sarkar, A. K. (2017). Application of neural network in modeling commuter choice behavior with a novel fuzzy access measure. In *Artificial Intelligence XXXIV: 37th SGAI International Conference on Artificial Intelligence, AI 2017, Cambridge, UK, December 12-14, 2017, Proceedings 37* (pp. 313-321). Springer International Publishing.

Mitchell, W. J. (1996). City of bits: space, place, and the infobahn. MIT press.

Mobaraki, A., & Oktay Vehbi, B. (2022). A Conceptual Model for Assessing the Relationship between Urban Morphology and Sustainable Urban Form. *Sustainability*, 14(5), 2884.

Mokhtari, K. E., Higdon, B. P., & Başar, A. (2019, November). Interpreting financial time series with SHAP values. In *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering* (pp. 166-172).

Morales, J., Flacke, J., & Zevenbergen, J. (2019). Modelling residential land values using geographic and geometric accessibility in Guatemala City. *Environment and Planning B: Urban Analytics and City Science*, *46*(4), 751-776.

Morris, J. M., Dumble, P. L., & Wigan, M. R. (1979). Accessibility indicators for transport planning. *Transportation Research Part A: General*, *13*(2), 91-109.

Mouratidis, K. (2018). Built environment and social well-being: How does urban form affect social life and personal relationships?. *Cities*, 74, 7-20.

Mouratidis, K. (2021). Urban planning and quality of life: A review of pathways linking the built environment to subjective well-being. *Cities*, 115, 103229.

Mouratidis, K., & Yiannakou, A. (2022). What makes cities livable? Determinants of neighborhood satisfaction and neighborhood happiness in different contexts. *Land use policy*, 112, 105855.

Moustafa, R. E. (2011). Parallel coordinate and parallel coordinate density plots. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(2), 134-148.

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., ... & Schwarz, N. (2013). Describing human decisions in agent-based models–ODD+ D, an extension of the ODD protocol. *Environmental Modelling & Software*, *48*, 37-48.

Municipal Corporation of Delhi. Final Report of fifth municipal valuation committee for municipal corporation of Delhi, 2022. Accessed from https://mcdonline.nic.in/portal/downloadFile/final_report_of_mvc_-v_committee_30.9.2022_230421104536436.pdf

Muñiz, I., Garcia-López, M. À., & Galindo, A. (2008). The effect of employment sub-centres on population density in Barcelona. *Urban Studies*, *45*(3), 627-649.

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences*, *116*(44), 22071-22080.

Musterd, S., Marcińczak, S., Van Ham, M., & Tammaru, T. (2017). Socioeconomic segregation in European capital cities. Increasing separation between poor and rich. *Urban geography*, 38(7), 1062-1083.

Muth, R. F. (1971). Cities and Housing: The Spatial Pattern of Urban Residential Land Use. Chicago:

Nabipour, M., Rosenberg, M. W., & Nasseri, S. H. (2022). The built environment, networks design, and safety features: An analysis of pedestrian commuting behavior in intermediate-sized cities. *Transport policy*, 129, 14-23.

Næss, P. (2022). Meta-analyses of built environment effects on travel: No new platinum standard. *Journal of Planning Education and Research*, 42(2), 199-205.

Næss, P., Strand, A., Wolday, F., & Stefansdottir, H. (2019). Residential location, commuting and non-work travel in two urban areas of different size and with different center structures. *Progress in Planning*, 128, 1-36.

Naikoo, M. W., Rihan, M., & Ishtiaque, M. (2020). Analyses of land use land cover (LULC) change and built-up expansion in the suburb of a metropolitan city: Spatio-temporal analysis of Delhi NCR using landsat datasets. *Journal of Urban Management*, 9(3), 347-359.

Nakamura, S., & Avner, P. (2021). Spatial distributions of job accessibility, housing rents, and poverty: The case of Nairobi. *J. Hous. Econ*., 51, 101743.

Nakshi, P., & Debnath, A. K. (2021). Impact of built environment on mode choice to major destinations in Dhaka. *Transportation research record*, 2675(4), 281-296.

Nasri, A., Zhang, L., & Younes, H. (2020). Analysis of the effect of multi-level urban form on bikeshare demand. *Journal of Transport and Land Use*, *13*(1), 389-408.

Nedovic-Budic, Z., Knaap, G. J., Shahumyan, H., Williams, B., & Slaev, A. (2016). Measuring urban form at community scale: Case study of Dublin, Ireland. *Cities*, 55, 148-164.

Negahban, A. (2017, December). Neural networks and agent-based diffusion models. In *2017 Winter Simulation Conference (WSC)* (pp. 1407-1418). IEEE.

Niemeier, D. A. (1997). Accessibility: an evaluation using consumer welfare. *Transportation*, *24*, 377-396.

Nieuwenhuijsen, M. J. (2020). Urban and transport planning pathways to carbon neutral, liveable and healthy cities; A review of the current evidence. *Environment international*, 140, 105661.

NITI Aayog (2022). Cities as Engines of Growth TA-9508: Strengthening the States for broad based urban development. Asian Development Bank Report. Accessed from https://www.niti.gov.in/sites/default/files/2022-05/Mod_CEOG_Executive_Summary_18052022.pdf

Niu, Q., Wang, Y., Xia, Y., Wu, H., & Tang, X. (2018). Detailed assessment of the spatial distribution of urban parks according to day and travel mode based on web mapping API: A case study of main parks in Wuhan. *International Journal of Environmental Research and Public Health*, *15*(8), 1725.

Nkeki, F. N., & Asikhia, M. O. (2019). Geographically weighted logistic regression approach to explore the spatial variability in travel behaviour and built environment interactions: Accounting simultaneously for demographic and socioeconomic characteristics. *Applied geography*, 108, 47-63.

Oke, J. B., Aboutaleb, Y. M., Akkinepally, A., Azevedo, C. L., Han, Y., Zegras, P. C., ... & Ben-Akiva, M. E. (2019). A novel global urban typology framework for sustainable mobility futures. *Environmental Research Letters*, 14(9), 095006.

Openshaw, S. (1984), The Modifiable Areal Unit Problem (Concepts and Techniques in Modern Geography: 38), Geo-Books, Norwich, UK.

OpenStreetMap contributors. (2023). OpenStreetMap: Delhi. Retrieved November 23, 2022, from https://www.openstreetmap.org

Özkazanç, S., & Sönmez, F. N. Ö. (2017). Spatial analysis of social exclusion from a transportation perspective: A case study of Ankara metropolitan area. *Cities, 67*, 74-84.

Öztürk, Ö., Gülgen, F., Bilgi, S., & Kiliç, B. (2018). Accessibility analysis of street networks using space syntax. In *7th International Conference on Cartography and GIS* (pp. 18-23).

Padeiro, M., Louro, A., & da Costa, N. M. (2019). Transit-oriented development and gentrification: a systematic review. *Transport Reviews*, 39(6), 733-754.

Páez A, & Wheeler DC. (2009) Geographically weighted regression. In International Encyclopedia of Human Geography. Elsevier, Oxford [UK]

Páez, A., Scott, D. M., & Morency, C. (2012). Measuring accessibility: positive and normative implementations of various accessibility indicators. *Journal of Transport Geography*, *25*, 141-153.

Palaiologou, G., Larimian, T., & Vaughan, L. (2021). The use of morphological description in neighbourhood planning: form-based assessment of physical character and design rules. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 14(4), 490-514.

Pang, H., & Zhang, M. (2019). Understand the multi-level effects of the built environment on trip-chaining behavior. *Transportation Research Record*, *2673*(4), 640-651.

Park, R. E. (1936). Human Ecology. American Journal of Sociology, 42(1), 1–15. http://www.jstor.org/stable/2768859

Patel, A., Crooks, A.T. & Koizumi, N. (2012). Slumulation: An Agent-based Modeling Approach to Slum Formations. *Journal of Artificial Societies and Social Simulation*, 15(4), 2. https://www.jasss.org/15/4/2.html.

Patel, E., & Kushwaha, D. S. (2020). Clustering cloud workloads: K-means vs gaussian mixture model. *Procedia Computer Science*, 171, 158-167.

Patnala, P. K., Parida, M., & Chalumuri, R. S. (2023). Gender differentials in travel behavior among TOD neighborhoods: Contributions of built environment and residential self-selection. *Travel Behaviour and Society*, 31, 333-348.

Paul, Davidoff, L., & Gold, N. N. (1970). Suburban action: advocate planning for an open society. Journal of the American Institute of Planners, 36(1), 12-21.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res*. 12, 2825.

Peimani, N., & Kamalipour, H. (2022). Assembling transit urban design in the global South: urban morphology in relation to forms of urbanity and informality in the public space surrounding transit stations. *Urban Science*, 6(1), 18.

Peng, Z. R. (1997). The jobs-housing balance and urban commuting. *Urban studies*, *34*(8), 1215-1235.

Pereira, M. F., Vale, D. S., & Santana, P. (2023). Is walkability equitably distributed across socio-economic groups?–A spatial analysis for Lisbon metropolitan area. *Journal of Transport Geography*, 106, 103491.

Perez, J., Fusco, G., Araldi, A., & Fuse, T. (2020). Identifying building typologies and their spatial patterns in the metropolitan areas of Marseille and Osaka. Asia-Pacific *Journal of Regional Science*, 4(1), 193-217.

Pilehforooshha, P., & Karimi, M. (2020). A local adaptive density-based algorithm for clustering polygonal buildings in urban block polygons. *Geocarto International*, 35(2), 141-167.

Pirie, G. H. (1979). Measuring accessibility: a review and proposal. *Environment and planning A*, *11*(3), 299-312.

Pirie, G. H. (1981). The possibility and potential of public policy on accessibility. *Transportation Research Part A: General*, *15*(5), 377-381.

Pirra, M., Kalakou, S., Carboni, A., Costa, M., Diana, M., & Lynce, A. R. (2021). A preliminary analysis on gender aspects in transport systems and mobility services: Presentation of a survey design. *Sustainability*, *13*(5), 2676.

Pozoukidou, G., & Chatziyiannaki, Z. (2021). 15-Minute City: Decomposing the new urban planning eutopia. *Sustainability*, 13(2), 928.

Pred, A. (1977). The choreography of existence: comments on Hägerstrand's time-geography and its usefulness. *Economic geography*, *53*(2), 207-221.

Quillian, L., & Lagrange, H. (2016). Socioeconomic segregation in large cities in France and the United States. Demography, 53(4), 1051-1084.

Radke, J., & Mu, L. (2000). Spatial decompositions, modeling and mapping service regions to predict access to social programs. *Geographic Information Sciences*, 6(2), 105-112.

Raghavendra, R. H. (2020). Literacy and health status of scheduled castes in India. *Contemporary Voice of Dalit*, 12(1), 97-110.

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, 48, 137-141.

Ramchandani, P., Paich, M., & Rao, A. (2017, August). Incorporating learning into decision making in agent based models. In *EPIA Conference on Artificial Intelligence* (pp. 789-800). Cham: Springer International Publishing.

Rana, D., Kumar, D., Kumari, M., & Kumari, R. (2022). Assessing the Impact of Delhi Metro Network Towards Urbanisation of Delhi-NCR. In *Geospatial Technology for Landscape and Environmental Management* (pp. 333-349). Springer, Singapore.

Rao, A. S., & Georgeff, M. P. (1997). Modeling rational agents within a BDI-architecture. *Readings in agents*, 317-328.

Rasca, S., & Saeed, N. (2022). Exploring the factors influencing the use of public transport by commuters living in networks of small cities and towns. *Travel Behaviour and Society*, 28, 249-263.

Raux, C., Ma, T-Y., Lemoy, R., & Ovtracht, N. (2014). A LUTI agent-based model of Lyon area: welfare analysis of some scenarios. *Transp Res Arena* (TRA)

Rawls, J. (1999). A Theory of Justice. Revised ed. Cambridge, MA: The Belknap Press of Harvard University Press.

Regasa, M. S., Nones, M., & Adeba, D. (2021). A review on land use and land cover change in Ethiopian basins. *Land*, 10(6), 585.

Rehman, A., & Jamil, F. (2021). Impact of urban residential location choice on housing, travel demands and associated costs: Comparative analysis with empirical evidence from Pakistan. *Transportation research interdisciplinary perspectives*, 10, 100357.

Röthlisberger, V., Zischg, A. P., & Keiler, M. (2017). Identifying spatial clusters of flood exposure to support decision making in risk management. *Science of the total environment, 598*, 593-603.

Roy, A. (2009). Why India cannot plan its cities: Informality, insurgence and the idiom of urbanization. Planning theory, 8(1), 76-87.

Ruo-Qi, L., & Jun-Hong, H. (2020, September). Prediction of housing price along the urban rail transit line based on GA-BP model and accessibility. In *2020 IEEE 5th International Conference on Intelligent Transportation Engineering (ICITE)* (pp. 487-492). IEEE.

Saberi, M., Wu, H., Amoh-Gyimah, R., Smith, J., & Arunachalam, D. (2017). Measuring housing and transportation affordability: A case study of Melbourne, Australia. Journal of transport geography, 65, 134-146.

Saghapour, T., & Moridpour, S. (2019). The role of neighbourhoods accessibility in residential mobility. *Cities*, *87*, 1-9.

Salem, M., Bose, A., Bashir, B., Basak, D., Roy, S., Chowdhury, I. R., & Tsurusaki, N. (2021). Urban expansion simulation based on various driving factors using a logistic regression model: Delhi as a case study. *Sustainability*, 13(19), 10805.

Sandoval, V., & Sarmiento, J. P. (2020). A neglected issue: Informal settlements, urban development, and disaster risk reduction in Latin America and the Caribbean. Disaster Prevention and Management: An International Journal, 29(5), 731-745.

Schirmer, P. M., & Axhausen, K. W. (2019). A multiscale clustering of the urban morphology for use in quantitative models. In *The Mathematics of Urban Morphology* (pp. 355-382). Birkhäuser, Cham.

Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., ... & Wijermans, N. (2017). A framework for mapping and comparing behavioural theories in models of social-ecological systems. *Ecological economics*, *131*, 21-35.

Schmidt, B. (2002). The modelling of human behavior: the PECS reference model. In: *Proceedings 14th European simulation symposium*, Dresden, Germany.

Schulze, J., Müller, B., Groeneveld, J., & Grimm, V. (2017). Agent-based modelling of social-ecological systems: achievements, challenges, and a way forward. *Journal of Artificial Societies and Social Simulation*, *20*(2).

Sharifi, A. (2019). Urban form resilience: A meso-scale analysis. *Cities*, 93, 238-252.

Sharifi, A., Roosta, M., & Javadpoor, M. (2021). Urban form resilience: A comparative analysis of traditional, semi-planned, and planned neighborhoods in Shiraz, Iran. *Urban Science*, 5(1), 18.

Sharifi, F., Nygaard, A., Stone, W. M., & Levin, I. (2021). Accessing green space in Melbourne: Measuring inequity and household mobility. *Landscape and Urban Planning*, 207, 104004.

Shen, Q. (1998). Location characteristics of inner-city neighborhoods and employment accessibility of low-wage workers. *Environment and planning B: Planning and Design*, *25*(3), 345-365.

Shen, Y., Sun, F., & Che, Y. (2017). Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the Central City of Shanghai. *Urban Forestry & Urban Greening*, 27, 59–68. https://doi.org/10.1016/j.ufug.2017.06.018

Shi, Z., Fonseca, J. A., & Schlueter, A. (2021). A parametric method using vernacular urban block typologies for investigating interactions between solar energy use and urban design. *Renewable Energy*, 165, 823-841.

Shin, E. J. (2020). Commuter benefits programs: Impacts on mode choice, VMT, and spillover effects. *Transport Policy*, 94, 11-22.

Shukla, N., Ma, J., Wickramasuriya, R., Huynh, N., & Perez, P. (2016). Modelling mode choice of individual in linked trips with artificial neural networks and fuzzy representation. *Artificial Neural Network Modelling*, 405-422.

Silva, C. A., Giannotti, M., & de Almeida, C. M. (2020). Dynamic modeling to support an integrated analysis among land use change, accessibility and gentrification. *Land use policy*, 99, 104992.

Silva, E. A. (2011). Cellular Automata and Agent Base Models for Urban Studies: From Pixels to Cells to Hexa-dpi's. *Urban remote sensing: Monitoring, synthesis and modeling in the urban environment*, 250, 323-334.

Silva, E. A., & Clarke, K. C. (2002). Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, environment and urban systems*, 26(6), 525-552.

Silva, E. A., & Clarke, K. C. (2005). Complexity, emergence and cellular urban models: lessons learned from applying SLEUTH to two Portuguese metropolitan areas. *European Planning Studies*, 13(1), 93-115.

Singh, B., Venkatramanan, V., & Deshmukh, B. (2022). Monitoring of land use land cover dynamics and prediction of urban growth using Land Change Modeler in Delhi and its environs, India. *Environmental Science and Pollution Research*, 29(47), 71534-71554.

So, K. S., Orazem, P. F., & Otto, D. M. (2001). The effects of housing prices, wages, and commuting time on joint residential and job location choices. *American journal of agricultural economics*, 83(4), 1036-1048.

Sodiq, A., Baloch, A. A., Khan, S. A., Sezer, N., Mahmoud, S., Jama, M., & Abdelaal, A. (2019). Towards modern sustainable cities: Review of sustainability principles and trends. *Journal of Cleaner Production*, 227, 972-1001.

Soja, E. (2009). The city and spatial justice. Justice spatiale/Spatial justice, 1(1), 1-5.

Soltani, A., Javadpoor, M., Shams, F., & Mehdizadeh, M. (2022). Street network morphology and active mobility to school: Applying space syntax methodology in Shiraz, Iran. *Journal of Transport & Health*, 27, 101493.

Song, Y., & Kim, H. (2018). Evolution of subway network systems, subway accessibility, and change of urban landscape: A longitudinal approach to Seoul Metropolitan Area. In *Intelligent Transportation and Planning: Breakthroughs in Research and Practice* (pp. 1087-1111). IGI Global.

State Election Commission Delhi. (2022). Ward delimitation data locality wise. Obtained from-https://sec.delhi.gov.in/sec/delimitation-2022-final

Stein, M., Nitschke, L., Trost, L., Dirschauer, A., & Deffner, J. (2022). Impacts of commuting practices on social sustainability and sustainable mobility. *Sustainability*, 14(8), 4469.

Stephen, Ezennia, I., & Hoskara, S. O. (2019). Methodological weaknesses in the measurement approaches and concept of housing affordability used in housing research: A qualitative study. *PloS one*, 14(8), e0221246.

Stokenberga, A. (2014). Does bus rapid transit influence urban land development and property values: a review of the literature. *Transport Reviews*, 34(3): 276-296.

Su, H., Han, G., Li, L., & Qin, H. (2021). The impact of macro-scale urban form on land surface temperature: An empirical study based on climate zone, urban size and industrial structure in China. *Sustainable Cities and Society*, 74, 103217.

Su, S., Zhang, H., Wang, M., Weng, M., & Kang, M. (2021). Transit-oriented development (TOD) typologies around metro station areas in urban China: A comparative analysis of five typical megacities for planning implications. *Journal of transport geography*, 90, 102939.

Sulemana, I., Nketiah-Amponsah, E., Codjoe, E. A., & Andoh, J. A. N. (2019). Urbanization and income inequality in Sub-Saharan Africa. *Sustainable cities and society*, 48, 101544.

Sun, B., & Yin, C. (2020). Impacts of a multi-scale built environment and its corresponding moderating effects on commute duration in China. *Urban Studies*, 57(10), 2115-2130.

Sun, B., Ermagun, A., & Dan, B. (2017). Built environmental impacts on commuting mode choice and distance: Evidence from Shanghai. *Transportation research part D: transport and environment*, 52, 441-453.

Sundaram, K. & Tendulker, S. D. (2003). Poverty among social and economic groups in India in 1990s. *Economic and Political Weekly*, 38(50), 5263–5276.

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018, April). Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP conference series: materials science and engineering* (Vol. 336, p. 012017). IOP Publishing.

Taaffe, E. J. (1996). Geography of transportation. Morton O'kelly.

Talen, E. (2022). Who can walk? An analysis of public amenity access in America's ten largest cities. *Environment and Planning B: Urban Analytics and City Science*, 23998083221142866.

Tammaru, T., Marcin´ Czak, S., Aunap, R., van Ham, M., & Janssen, H. (2020). Relationship between income inequality and residential segregation of socioeconomic groups. *Regional Studies*, 54(4), 450-461.

Tang, B. S., Wong, K. K., Tang, K. S., & Wai Wong, S. (2021). Walking accessibility to neighbourhood open space in a multi-level urban environment of Hong Kong. *Environment and Planning B: Urban Analytics and City Science*, 48(5), 1340-1356.

Tanwanichkul, L., Kaewwichian, P., & Pitaksringkarn, J. (2019). Car ownership demand modeling using machine learning: Decision trees and neural networks. *GEOMATE Journal*, *17*(62), 219-230.

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*(1), 1-16.

Tao, Z., Yao, Z., Kong, H., Duan, F., & Li, G. (2018). Spatial accessibility to healthcare services in Shenzhen, China: Improving the multi-modal two-step floating catchment area method by estimating travel time via online map APIs. *BMC health services research*, *18*(1), 1-10.

Tayal, D., & Mehta, A. K. (2021). Working women, Delhi Metro and Covid-19: a case study in Delhi-NCR. The Indian Journal of Labour Economics, 64(2), 389-413.

Tiwari, C., Bhattacharjee, S., Sethi, P., & Chakrabarti, D. (2022). Internal Migration and Rural Inequalities in India. *Population Research and Policy Review*, 41(4), 1673-1698.

Tobler, W. (1979) Cellular geography, in Philosophy and Geography (eds S. Gale and G. Olosson), D. Reidel, Dordrecht, pp. 279–386.

Tollefsen, A. F., Strand, H., & Buhaug, H. (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49(2), 363-374.

Tomasiello, D. B., Giannotti, M., & Feitosa, F. F. (2020). ACCESS: An agent-based model to explore job accessibility inequalities. *Computers, Environment and Urban Systems*, *81*, 101462.

Tripathi, S. (2021). How does urbanization affect the human development index? A cross-country analysis. *Asia-Pacific Journal of Regional Science*, 5(3), 1053-1080.

Türkoğlu, H., Terzi, F., Salihoğlu, T., Bölen, F., & Okumuş, G. (2019). Residential satisfaction in formal and informal neighborhoods: the case of Istanbul, Turkey. *Archnet-IJAR: International Journal of Architectural Research*, *13*(1), 112-132.

U.K Labour Force Survey. (2018). Office for National Statistics, Government of UK.

U.S Census Bureau. (2019). American Community Surveys.

U.S. Environmental Protection Agency. (2017). Smart Growth Zoning Codes. Retrieved from [https://www.epa.gov/sites/default/files/2017-06/documents/sm_growth_guide.pdf]

Ulam, S. (1976). Adventures of a Mathematician (Charles Scribner's Sons, New York)

UNDESA: United Nations Department of Economic and Social Affairs, (2018); Population Division

United Nations, Department of Economic and Social Affairs, Population Division (2018). World Urbanization Prospects: The 2018 Revision, Online Edition

United Nations. (2014, October 29). Liveable Cities Key to Sustainable Development, Secretary-General Stresses in Message for International Day [Press Release]. https://www.un.org/press/en/2014/sgsm16293.doc.htm

Vale, D. S., & Pereira, M. (2017). The influence of the impedance function on gravity-based pedestrian accessibility measures: A comparative analysis. *Environment and Planning B: Urban Analytics and City Science*, *44*(4), 740-763.

Vale, D. S., Saraiva, M., & Pereira, M. (2016). Active accessibility: A review of operational measures of walking and cycling accessibility. *Journal of transport and land use*, *9*(1), 209-235.

Van der Hoog, S. (2018). Surrogate Modelling in (and of) Agent-Based Models: A Prospectus. *Computational Economics*, 53(3): 1245-1263.

Van der Vlugt, A. L., Curl, A., & Scheiner, J. (2022). The influence of travel attitudes on perceived walking accessibility and walking behaviour. *Travel behaviour and society*, 27, 47-56.

Van Heerden, Q., Karsten, C., Holloway, J., Petzer, E., Burger, P., & Mans, G. (2022). Accessibility, affordability, and equity in long-term spatial planning: Perspectives from a developing country. *Transport Policy*, 120, 104-119.

van Herick, D., & Mokhtarian, P. L. (2020). How much does the method matter? An empirical comparison of ways to quantify the influence of residential self-selection. *Travel Behaviour and Society*, 18, 68-82.

Van Wee, B., & Cao, X. J. (2022). Residential self-selection in the relationship between the built environment and travel behavior: A literature review and research agenda. *Advances in Transport Policy and Planning*, 9, 75-94.

Van Wee, B., De Vos, J., & Maat, K. (2019). Impacts of the built environment and travel behaviour on attitudes: Theories underpinning the reverse causality hypothesis. *Journal of transport geography*, 80, 102540.

Vasconcelos, A. S., & Farias, T. L. (2012). Evaluation of urban accessibility indicators based on internal and external environmental costs. *Transportation Research Part D: Transport and Environment*, *17*(6), 433-441.

Verburg, P. H., Crossman, N., Ellis, E. C., Heinimann, A., Hostert, P., Mertz, O., ... & Zhen, L. (2015). Land system science and sustainable development of the earth system: A global land project perspective. *Anthropocene*, *12*, 29-41.

Vickerman, R. W. (1974). Accessibility, Attraction, and Potential: A Review of Some Concepts and Their Use in Determining Mobility. *Environment and Planning A: Economy and Space*, 6(6): 675–691.

Viegas, J. M., Martinez, L. M., & Silva, E. A. (2009). Effects of the modifiable areal unit problem on the delineation of traffic analysis zones. *Environment and Planning B: Planning and Design*, 36(4), 625-643.

Viggers, H., & Howden-Chapman, P. (2011). Urban form, public transport and mortgagee sales. *Growth misconduct? Avoiding sprawl and improving urban intensification in New Zealand*, 159-177.

Vogiazides, L., & Mondani, H. (2023). Neighbourhood trajectories in Stockholm: Investigating the role of mobility and in situ change. *Applied geography*, 150, 102823.

Von Thunen, J. H. (1826) Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalokonomie (Rostock)

Wachs, M., & Kumagai, T. G. (1973). Physical accessibility as a social indicator. *Socio-Economic Planning Sciences*, *7*(5), 437-456.

Waddell, P. (2000). A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim. *Environment and Planning B: Planning and Design*, 27(2): 247 – 263.

Waddell, P., Borning, A., Noth, M., Freier, N., Becke, M., & Ulfarsson, G. (2003). Microsimulation of urban development and location choices: Design and implementation of UrbanSim. *Networks and spatial economics*, 3(1), 43-67.

Wan, N., Zou, B., & Sternberg, T. (2012). A three-step floating catchment area method for analyzing spatial access to health services. *International Journal of Geographical Information Science*, *26*(6), 1073-1089.

Wang, D., Cao, X., & Huang, X. (2021). Equity of accessibility to health care services and identification of underserved areas. *Chinese Geographical Science*, 31(1), 167-180.

Wang, F. (2012). Measurement, optimization, and impact of health care accessibility: a methodological review. *Annals of the Association of American Geographers*, 102(5): 1104-1112.

Wang, J., & Biljecki, F. (2022). Unsupervised machine learning in urban studies: A systematic review of applications. *Cities*, 129, 103925.

Wang, J., Du, F., Huang, J., & Liu, Y. (2020). Access to hospitals: Potential vs. observed. *Cities*, *100*, 102671.

Wang, J., Wu, J., & Che, Y. (2019). Agent and system dynamics-based hybrid modeling and simulation for multilateral bidding in electricity market. *Energy*, *180*, 444-456.

Wang, S. W., Munkhnasan, L., & Lee, W. K. (2021). Land use and land cover change detection and prediction in Bhutan's high altitude city of Thimphu, using cellular automata and Markov chain. *Environmental Challenges*, 2, 100017.

Wang, Y., Chen, B. Y., Yuan, H., Wang, D., Lam, W. H., & Li, Q. (2018). Measuring temporal variation of location-based accessibility using space-time utility perspective. *Journal of Transport Geography*, *73*, 13-24.

Wang, Y., Li, Y., Huang, Y., Yi, C., & Ren, J. (2020). Housing wealth inequality in China: An urban-rural comparison. *Cities*, 96, 102428.

Wang, Z., Yang, Y., Yao, E., Zhu, Y., & Rao, Z. (2022). An analysis of commute mode choice behavior considering the impacts of built environment in Beijing. *Transportation letters*, 14(7), 733-739.

Weber, J., & Kwan, M. P. (2002). Bringing time back in: A study on the influence of travel time variations and facility opening hours on individual accessibility. *The professional geographer*, *54*(2), 226-240.

Weibull, J. W. (1976). An axiomatic approach to the measurement of accessibility. *Regional Science and Urban Economics*, 6(4): 357–379.

Weibull, J. W. (1980). On the Numerical Measurement of Accessibility. *Environment and Planning A*, 12(1): 53–67.

Weld, G., Jang, E., Li, A., Zeng, A., Heimerl, K., & Froehlich, J. E. (2019, October). Deep learning for automatically detecting sidewalk accessibility problems using streetscape imagery. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 196-209).

Wheaton, W. C. (1976). On the optimal distribution of income among cities. *Journal of Urban economics*, 3(1), 31-44.

Wheaton, WC. (1974). A comparative static analysis of urban spatial structure. *J Econ Theory*, 9(2), 223–237

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*, 7(2), 161-187.

White, M. J. (1999). Urban Areas with Decentralized Employment: Theory and Empirical Work, in Chesire, P. and E.S. Mills (Eds.), Handbook of Regional and Urban Economics, Volume 3 Applied Urban Economics, North-Holland.

Wickstrom, G. B. (1971). Defining balanced transportation: a question of opportunity. *Traffic Quarterly*, 25: 337–350

Wilensky, U. (1999). NetLogo. In Center for Connected Learning and Computer-Based Modeling. Evanston, IL: Northwestern University. http://ccl.northwestern.edu/netlogo/.

Wilson, A. G. (1971). A Family of Spatial Interaction Models, and Associated Developments. *Environment and Planning A: Economy and Space*, 3(1): 1–32.

Wolday, F. (2018). Built environment and car driving distance in a small city context. *Journal of Transport and Land Use*, 11(1), 747-767.

Wolday, F. (2023). The effect of neighbourhood and urban center structures on active travel in small cities. *Cities*, 132, 104050.

Wolday, F., Næss, P., & Cao, X. J. (2019). Travel-based residential self-selection: A qualitatively improved understanding from Norway. *Cities*, 87, 87-102.

Woodcock, J., Banister, D., Edwards, P., Prentice, A. M., & Roberts, I. (2007). Energy and transport. *The Lancet*, *370*(9592), 1078-1088.

Wu, C., Chung, H., Liu, Z., & Kim, I. (2021). Examining the effects of the built environment on topological properties of the bike-sharing network in Suzhou, China. *International Journal of Sustainable Transportation*, *15*(5), 338-350.

Wu, H., Jiao, H., Yu, Y., Li, Z., Peng, Z., Liu, L., & Zeng, Z. (2018). Influence factors and regression model of urban housing prices based on internet open access data. *Sustainability*, *10*(5), 1676.

Wu, J., He, Q., Chen, Y., Lin, J., & Wang, S. (2020). Dismantling the fence for social justice? Evidence based on the inequity of urban green space accessibility in the central urban area of Beijing. *Environment and Planning B: Urban Analytics and City Science*, 47(4), 626-644.

Wu, N., & Silva, E. A. (2010). Artificial intelligence solutions for urban land dynamics: a review. Journal of Planning Literature, 24(3), 246-265.

Wu, Y., Mashhoodi, B., Patuano, A., Lenzholzer, S., Zertuche, L. N., & Acred, A. (2022). Heat-prone neighbourhood typologies of European cities with temperate climate. *Sustainable Cities and Society*, 87, 104174.

Xu, G., Dong, T., Cobbinah, P. B., Jiao, L., Sumari, N. S., Chai, B., & Liu, Y. (2019). Urban expansion and form changes across African cities with a global outlook: Spatiotemporal analysis of urban land densities. *Journal of Cleaner Production*, 224, 802-810.

Xu, S., Li, Z., Zhang, C., Huang, Z., Tian, J., Luo, Y., & Du, H. (2021). A method of calculating urban-scale solar potential by evaluating and quantifying the relationship between urban block typology and occlusion coefficient: A case study of Wuhan in Central China. *Sustainable Cities and Society*, 64, 102451.

Xu, T., Gao, J., Coco, G., & Wang, S. (2020). Urban expansion in Auckland, New Zealand: A GIS simulation via an intelligent self-adapting multiscale agent-based model. International Journal of Geographical Information Science, 34(11), 2136-2159.

Xu, Y., & Heikkila, E. J. (2020). How can cities learn from each other? Evidence from China's five-year plans. *Journal of Urban Management*, 9(2), 216-227.

Xue, H., Cheng, X., Jia, P., & Wang, Y. (2020). Road network intersection density and childhood obesity risk in the US: a national longitudinal study. *Public health*, 178, 31-37.

Yamu, C., Van Nes, A., & Garau, C. (2021). Bill Hillier's legacy: Space syntax—A synopsis of basic concepts, measures, and empirical application. *Sustainability*, *13*(6), 3394.

Yan, X. (2020). Evaluating household residential preferences for walkability and accessibility across three U.S. regions. *Transportation Research Part D: Transport and Environment*, 80: 102255.

Yan, X., Zhou, J., Sheng, F., & Niu, Q. (2022). Influences of Built Environment at Residential and Work Locations on Commuting Distance: Evidence from Wuhan, China. *ISPRS International Journal of Geo-information*, 11(2), 124. https://doi.org/10.3390/ijgi11020124

Yang, L., Ding, C., Ju, Y., & Yu, B. (2021). Driving as a commuting travel mode choice of car owners in urban China: Roles of the built environment. *Cities*, 112, 103114.

Yang, W., Yang, R., & Zhou, S. (2022). The spatial heterogeneity of urban green space inequity from a perspective of the vulnerable: A case study of Guangzhou, China. *Cities*, 130, 103855.

Yen, T. W., Filippov, M., & Cheong, S. A. (2019). Emergent urban morphologies in an agent-based model of housing. In *Network Theory and Agent-Based Modeling in Economics and Finance* (pp. 227-246). Springer, Singapore.

Yu, H., & Peng, Z. R. (2019). Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted Poisson regression. *Journal of Transport Geography*, 75, 147-163.

Yu, S., Shang, C., Yu, Y., Zhang, S., & Yu, W. (2016). Prediction of bus passenger trip flow based on artificial neural network. *Advances in Mechanical Engineering*, *8*(10), 1687814016675999.

Yuan, C., & Yang, H. (2019). Research on K-value selection method of K-means clustering algorithm. *J*, 2(2), 226-235.

Yuan, Y., Wang, M., Zhu, Y., Huang, X., & Xiong, X. (2020). Urbanization's effects on the urban-rural income gap in China: A meta-regression analysis. *Land Use Policy*, 99, 104995.

Zang, P., Lu, Y., Ma, J., Xie, B., Wang, R., & Liu, Y. (2019). Disentangling residential self-selection from impacts of built environment characteristics on travel behaviors for older adults. *Social Science & Medicine*, 238, 112515.

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2018). An object-based convolutional neural network (OCNN) for urban land use classification. *Remote sensing of environment*, *216*, 57-70.

Zhang, H., & Vorobeychik, Y. (2019). Empirically grounded agent-based models of innovation diffusion: a critical review. *Artificial Intelligence Review*, *52*, 707-741.

Zhang, H., Jin, X., Wang, L., Zhou, Y., & Shu, B. (2015). Multi-agent based modeling of spatiotemporal dynamical urban growth in developing countries: simulating future scenarios of Lianyungang city, China. *Stochastic environmental research and risk assessment*, *29*, 63-78.

Zhang, J., Xu, L., Shabunko, V., Tay, S. E. R., Sun, H., Lau, S. S. Y., & Reindl, T. (2019). Impact of urban block typology on building solar potential and energy use efficiency in tropical high-density city. *Applied Energy*, 240, 513-533.

Zhang, L., Chen, P., & Hui, F. (2022). Refining the accessibility evaluation of urban green spaces with multiple sources of mobility data: A case study in Shenzhen, China. *Urban Forestry & Urban Greening*, 70, 127550.

Zhang, L., Ye, Y., Zeng, W., & Chiaradia, A. (2019). A systematic measurement of street quality through multi-sourced urban data: A human-oriented analysis. *International journal of environmental research and public health*, *16*(10), 1782.

Zhang, Y., Dai, X., Yu, X., & Gao, N. (2020). Urban integration of land-deprived households in China: Quality of living and social welfare. *Land Use Policy*, 96, 104671.

Zhang, Y., Li, Z., & Zhang, Y. (2020). Validation and calibration of an agent-based model: A surrogate approach. *Discrete Dynamics in Nature and Society*, *2020*.

Zhao, C., Nielsen, T. A. S., Olafsson, A. S., Carstensen, T. A., & Meng, X. (2018). Urban form, demographic and socio-economic correlates of walking, cycling, and e-biking: Evidence from eight neighborhoods in Beijing. *Transport Policy*, 64, 102-112.

Zhao, P., Li, S., & Liu, D. (2020). Unequable spatial accessibility to hospitals in developing megacities: New evidence from Beijing. *Health & Place*, 65, 102406.

Zheng, B., Cai, Y., Zeng, F., Lin, M., Zheng, J., Chen, W., ... & Guo, Y. (2021). An interpretable model-based prediction of severity and crucial factors in patients with COVID-19. *BioMed Research International*, 2021.

Zheng, Z., Xia, H., Ambinakudige, S., Qin, Y., Li, Y., Xie, Z., ... & Gu, H. (2019). Spatial accessibility to hospitals based on web mapping API: An empirical study in Kaifeng, China. *Sustainability*, *11*(4), 1160.

Zhou, C., Li, S., & Wang, S. (2018). Examining the impacts of urban form on air pollution in developing countries: A case study of China’s megacities. *International journal of environmental research and public health*, 15(8), 1565.

Zhou, L., Dang, X., Sun, Q., & Wang, S. (2020). Multi-scenario simulation of urban land change in Shanghai by random forest and CA-Markov model. *Sustainable Cities and Society*, 55, 102045.

Zhu, P., Ho, S. N., Jiang, Y., & Tan, X. (2020). Built environment, commuting behaviour and job accessibility in a rail-based dense urban context. *Transportation Research Part D: Transport and Environment*, 87, 102438.

Zhu, P., Zhao, S., & Jiang, Y. (2022). Residential segregation, built environment and commuting outcomes: Experience from contemporary China. *Transport Policy*, 116, 269-277.

Zhu, Y., & Diao, M. (2020). Crowdsourcing-data-based dynamic measures of accessibility to business establishments and individual destination choices. *Transportation Research Part D: Transport and Environment*, *87*, 102382.

Zhuge, C., Shao, C., Gao, J., Dong, C., & Zhang, H. (2016). Agent-based joint model of residential location choice and real estate price for land use and transport model. *Computers, environment and urban systems*, *57*, 93-105.

Zuo, Y., Fu, X., Liu, Z., & Huang, D. (2021). Short-term forecasts on individual accessibility in bus system based on neural network model. *Journal of Transport Geography*, *93*, 103075.

Appendix A

Household Survey

1 Demographic Details

1.1 Name of Respondent (Can be left blank)

1.2 Gender

1.3 Age

1.4 Marital Status

1.5 Education

1.6 Household Members

1.7 Does households have any school going children

2 Socio-Economic Characteristics

2.1 Number of family members who go for work.

2.2 Monthly Household Income

2.3 Caste

2.4 Type of Job:
(Daily wage earner/ Street Vendor/ Own Business/ Private salaried/ Government Servant Other)

2.5 Profession

2.6 Type and Number of vehicles owned.
(i) Cars -
(ii) Two-wheelers –

3. Household Characteristics

3.1 Type of Dwelling – (Independent house/ Flat/ Shanty)

3.2 Number of Floors

3.3. Household Ownership

3.4 Average land price for one unit area in this lane

3.5 House Base Floor Area and unit

3.6 Monthly rental (if applicable)

3.7 Did you relocate in last 50 years?

3.8 If yes, previous location area

3.9 Reason of relocation –

i. To live near the workplace
ii. To live near amenities such as city center, transit stations, schools, hospital, parks, etc.
iii. To live in big house
iv. To live in clean and pollution free environment.

v. To live in less dense area
vi. To live with people of my caste group
vii. To lower the house rent
viii. Marriage
ix. To buy my own house
x. NA
xi. Other

3.10 Year of relocation

4 Travel Characteristics

4.1 Job location

4.2 Distance to job location

4.3(a) Current Mode of travel from home to workplace

4.3(b) Current Mode of travel from workplace to home

4.4 Total monthly spending on your travel to workplace

4.5 Reason for not using metro / bus: (Choose NA only if travel mode is walk, bus or metro)

i. Metro/Bus station is not near to my home or workplace
ii. Transit time in metro/bus is higher
iii. No direct route so it takes more time as I need to change metro/bus
iv. I need to travel at different locations in a day
v. Health reasons
vi. Cab service from office
vii. Metro/Bus stations are over-crowded / long queue
viii. Joint travel in car with family members or friends/colleagues Job location is nearby
ix. NA
x. Privacy issues
xi. Other

4.6 Rank the given three attributes of your mode of travel in decreasing order (Rank 1= Very important, Rank 3 = Not so important)

Travel Attributes	Rank
Travel Expenditure	
Travel Time	
Travel Comfort	

4.7 Nature of your trip *

i. I travel alone throughout from home to workplace
ii. I share part of my trip with my spouse / family member
iii. I share part of my trip with my colleagues / friends / neighbors I share my trip completely with my my spouse/ family member
iv. I share my trip completely with my colleagues / friends /neighbors

4.8(a) Have you changed your mode of travel to job in last 10 year?

4.9(b) If yes, mention the year when there was change in mode of travel

4.8(c) What was their earlier mode of travel?

4.8(d) Reason for the latest change in travel mode, if any?

i. Increase in income
ii. Availability of Metro from home to job location
iii. Availability of bus from home to job location
iv. Safety or health issues
v. To reduce travel time
vi. To reduce travel cost
vii. To make joint travel
viii. NA
ix. Other:

4.9 Travel frequency to job destination in a day

4.10 Distance to metro station/ bus stand from current job location (kms)

4.11 Availability of authorized parking area at workplace

4.12(a) Time at which you leave home for your workplace usually

4.12(b) Time at which you leave workplace for your home usually

4.13 Time spent in travelling to job from home (in minutes)

4.14 Since you are working in Delhi, your travel time to job has –
(Increased / Decreased / Almost same)

4.15(a) If your travel time to job has increased, this is due to (Choose NA otherwise) *

i. I have relocated further away from the job location
ii. I have changed my job which is further away from my residence Increase in traffic jams
iii. Long waiting time at transit stations NA
iv. Other:

4.15(b) If your travel time to job has decreased, this is due to (Choose NA otherwise) *

i. Relocated near to job destination
ii. Taken up a job closer to my residence
iii. Better route coverage by metro/ buses
iv. Use of high speed personal vehicle
v. Improved Road infrastructure such as new bridges
vi. NA
vii. Other:

5 Neighbourhood Characteristics

5.1 If given a chance to relocate, Rank your preference to the following Neighborhood location choices (1 = highest rank, 8 = lowest rank) *

Neighbourhood location	Score
Proximity to workplace	
Proximity to public amenities like parks, health centres, schools.	

Proximity to metro or bus stations	
Proximity to major road	
Proximity to market	
Proximity to religious center like temple, mosque	
Proximity to relatives or people of same community	

5.2 For the following amenities mark the travel MODE and travel FREQUENCY *

(Frequency - Everyday, Once or twice a week, Sometimes in a month; Rarely;
Mode - Car / Bike; Metro / Bus; Walk / Cycle; Rickshaw/Tempo/Tuk-Tuk; Auto/Cab)

Amenities	Travel Mode	Travel Frequency
Market		
Metro Station		
Bus Stop		
District Centre		

5.3 Distance to following amenities

Amenities	0-1 Km	1-2 Km	2-5 Km	More than 5 kms
Market				
Metro				
Bus Stop				
Health Centre				
Park				

5.4 Rate level of cleanliness in your house lane (1 – very unhygienic; 5 – very clean)

5.5 Rate Level of Safety in your house lane (1 – very safe; 5 – very safe)

5.6 Ease of walking in your house lane

Street parameters	Yes/No
Availability of footpaths/pedestrian ways	
Presence of potholes	
Roadside greenery	
Does the road has slope	
Does the road has frequentt jams	

5.7 Street width

5.8 Neighbourhood population density in house lane (High, Low, Medium)

5.9 Distance from Connaught place