The Urban Book Series

Robert Goodspeed
Raja Sengupta
Marketta Kyttä
Christopher Pettit *Editors*

Intelligence for Future Cities

Planning Through Big Data and Urban Analytics

Springer

The Urban Book Series

The Urban Book Series is a resource for urban studies and geography research worldwide. It provides a unique and innovative resource for the latest developments in the field, nurturing a comprehensive and encompassing publication venue for urban studies, urban geography, planning and regional development.

The series publishes peer-reviewed volumes related to urbanization, sustainability, urban environments, sustainable urbanism, governance, globalization, urban and sustainable development, spatial and area studies, urban management, transport systems, urban infrastructure, urban dynamics, green cities and urban landscapes. It also invites research which documents urbanization processes and urban dynamics on a national, regional and local level, welcoming case studies, as well as comparative and applied research.

The series will appeal to urbanists, geographers, planners, engineers, architects, policy makers, and to all of those interested in a wide-ranging overview of contemporary urban studies and innovations in the field. It accepts monographs, edited volumes and textbooks.

Indexed by Scopus.

Robert Goodspeed · Raja Sengupta ·
Marketta Kyttä · Christopher Pettit
Editors

Intelligence for Future Cities

Planning Through Big Data and Urban Analytics

Springer

Editors
Robert Goodspeed
Taubman College of Architecture
and Urban Planning
University of Michigan
Ann Arbor, MI, USA

Marketta Kyttä
Department of Built Environment
Aalto University
Espoo, Finland

Raja Sengupta
Department of Geography
McGill University
Montreal, QC, Canada

Christopher Pettit
City Futures Research Centre
University of New South Wales
Sydney, NSW, Australia

ISSN 2365-757X ISSN 2365-7588 (electronic)
The Urban Book Series
ISBN 978-3-031-31745-3 ISBN 978-3-031-31746-0 (eBook)
https://doi.org/10.1007/978-3-031-31746-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The international conference Computational Urban Planning and Urban Management (CUPUM) is one of the premier international conferences for the exchange of ideas and applications of computer technologies to address a range of social and environmental problems relating to urban areas. The first conference took place in 1989 in Hong Kong. Since then, this biennial conference has been hosted in cities across Asia, Australia, Europe, North America, and South America (Table 1).

Table 1 Past CUPUM conferences

Number	Year	Place	Country
I	1989	Hong Kong	China
II	1991	Oxford	UK
III	1993	Atlanta	USA
IV	1995	Melbourne	Australia
V	1997	Mumbai	India
VI	1999	Venice	Italy
VII	2001	Honolulu	USA
VIII	2003	Sendai	Japan
IX	2005	London	UK
X	2007	Iguazu Falls	Brazil
XI	2009	Hong Kong	China
XII	2011	Lake Louise (Calgary/Banff)	Canada
XIII	2013	Utrecht	The Netherlands
XIV	2015	Boston	USA
XV	2017	Adelaide	Australia
XVI	2019	Wuhan	China
XVII	2021	Helsinki	Finland
XVIII	2023	Montreal	Canada

Table 2 Board of Directors of CUPUM

Name	Region	Institution	Country	E-mail
Chris Pettit (Chair)	Australia	University of New South Wales	AUS	c.pettit@unsw.edu.au
Sharon Biermann	Australia	University of Western Australia	AUS	sharon.biermann@uwa.edu.au
Adam Dennett	Europe	University College London	GBR	a.dennett@ucl.ac.uk
Joseph Ferreira	North America	Massachusetts Institute of Technology	USA	jf@mit.edu
Robert Goodspeed	North America	University of Michigan	USA	rgoodspe@umich.edu
Marketta Kyttä	Europe	Aalto University	FIN	marketta.kytta@aalto.fi
Wenfeng Li	Asia	University of Hong Kong	CHN	wfli@hku.hk
Renee Sieber	North America	McGill University	CAN	renee.sieber@mcgill.ca
Atsushi Suzuki	Asia	Meijo University	JPN	atsuzuki@meijo-u.ac.jp
Qingming Zhan	Asia	Wuhan University	CHN	qmzhan@whu.edu.cn

This volume is being published in conjunction with the 18th CUPUM conference, to be hosted by McGill University in Montreal, Canada, in June 2023.

This volume was led by the editors, who are also members of the CUPUM board (Table 2) and the CUPUM 2023 Organizing Committee, comprising Renée Sieber, Raja Sengumpta, David Wachsmuth, and Suthee "Peck" Sangiabut from McGill University; Xinyue Ye from Texas A&M University; and Robert Goodspeed from the University of Michigan. The editors produced the volume with support from these two groups, as well as the CUPUM advisers to the board (Table 3). After an initial call for submissions, the book editors received 41 draft book chapters, 16 of which appear in the final volume. All published chapters underwent a double-blind peer-review process and will be presented at the CUPUM 2023 conference in Montreal, June 20–22, 2023. Many of the submitted chapters that were not included in this volume are to be presented as short papers and are likely to appear elsewhere in the scholarly literature. After a virtual 2021 conference during the COVID-19 pandemic, we hope this volume and a lively in-person event will foster the vibrant intellectual community that is CUPUM.

As editors, we wish to thank the many people who made this volume possible. The authors submitted very high-quality chapters, and we also thank them for

Table 3 Advisers to the CUPUM board

Name	Institution	Country	E-mail
Michael Batty (Chair)	University College London	GBR	m.batty@ucl.ac.uk
Karl Kim	University of Hawaii	USA	karlk@hawaii.edu
Richard Klosterman	University of Akron	USA	dick.klosterman@gmail.com
Kazuaki Miyamoto	Tokyo City University	JPN	miyamoto@tcu.ac.jp
Paola Rizzi	Università degli Studi di Sassari	BRA	rizzi@uniss.it
John Stillwell	University of Leeds	GBR	j.c.h.stillwell@leeds.ac.uk
Anthony G. O. Yeh	University of Hong Kong	CHN	hdxugoy@hkucc.hku.hk
Ray Wyatt	University of Melbourne	AUS	raywyatt3@gmail.com
Michael Wegener	University of Dortmund	DEU	mw@spiekermann-wegener.de
Antônio Nélson Rodrigues da Silva	University of Sao Paulo	BRA	anelson@sc.usp.br
Stan Geertman	Utrecht University	NLD	s.c.m.geertman@uu.nl
Andrew Allan	University of South Australia	AUS	andrew.allan@unisa.edu.au

their diligent and careful revisions in response to feedback from the reviewers and editors. All of the peer reviewers for the book, which included members of the CUPUM boards as well as other scholars involved in the CUPUM community, provided invaluable feedback to strengthen the overall quality. In addition, Springer has remained a supportive publishing partner, and this marks the sixth volume published in conjunction with a CUPUM event. In particular, we thank Juliana Pitanguy, the editor of the *Urban Book Series*, for her support for this project, and Sanjiev Kumar for professionally leading the final production. The other volumes published in conjunction with prior CUPUM conferences are *Planning Support Systems for Sustainable Urban Development* (2013), *Planning Support Systems and Smart Cities* (2015), *Planning Support Science for Smarter Urban Futures* (2017), *Computational Urban Planning and Management for Smart Cities* (2019), and *Urban Informatics and Future Cities* (2021).

After the editors issued the call for submissions, we waited with some trepidation to see what response we would receive, given the impact of the pandemic on international travel and scholarly productivity in general. However, by the deadline we had received many fascinating and innovative submissions, which showed that the CUPUM community was thriving and also tackling some of the most pressing urban problems in new and different ways. Although the many challenges facing our urban communities are daunting, we are inspired by the creativity and passion of the

CUPUM community to focus on how computational methods and techniques can be applied to address these challenges.

Ann Arbor, MI, USA
2023

Robert Goodspeed
Raja Sengupta
Marketta Kyttä
Christopher Pettit

Contents

Chapter 1
Introduction

Robert Goodspeed, Raja Sengupta, Marketta Kyttä, and Christopher Pettit

Abstract This introductory chapter provides a summary of the book's contents. The book is organized into three parts: Digital Cities, Mobility Futures, and Fine-scale Urban Analysis. The chapters contain innovative research about smart cities, urban platforms, bikeability, ride-hailing, walkability, planning support systems, urban heat mitigation, and urban growth modeling. The 16 chapters are to be presented at the 2023 Computational Urban Planning and Urban Management conference at McGill University in June.

Keywords Urban planning · Geographic information systems · Planning support systems · Urban analytics

This book contains chapters to be presented at the 18th Computational Urban Planning and Urban Management (CUPUM) conference at McGill University in June 2023. Similar volumes have been published biennially in conjunction with this conference since 2013. CUPUM comprises a global community of scholars conducting research at the frontiers of computational approaches to urban problems, in fields including urban modeling, urban analytics, geospatial analysis, and planning support systems. The volume provides examples of innovative research in several

R. Goodspeed (✉)
Taubman College of Architecture and Urban Planning, University of Michigan, Ann Arbor, Michigan, United States
e-mail: rgoodspe@umich.edu

R. Sengupta
Department of Geography, McGill University, Montreal, Canada
e-mail: raja.sengupta@mcgill.ca

M. Kyttä
Aalto University, Espoo, Finland
e-mail: marketta.kytta@aalto.fi

C. Pettit
University of New South Wales, Kensington, Australia
e-mail: c.pettit@unsw.edu.au

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_1

scholarly areas, as well as a bellwether for key trends in the field. Whereas the prior volume was prepared as the world exited the coronavirus disease 2019 (COVID-19) global pandemic, life has largely returned to normal in most parts of the world. However, the effects of the pandemic will continue to be seen for years to come and are evident within the contributed chapters. Furthermore, topics of longstanding interest such as walkability, travel behavior, planning support systems, smart cities, and many more make appearances in these pages. The overview describes the topics of the chapters and traces out how the volume's chapters speak to global urban trends and issues just emerging on the horizon.

To create the volume, the editors issued a call for submissions, which resulted in 41 draft chapters submitted in fall 2022. After our editorial process, which included double-blind peer review of all chapters, we identified 16 of the best chapters for inclusion in this book. This volume's chapters are to be presented at the CUPUM conference in June 2023, hosted by McGill University in Quebec, Canada. This work is somewhat shorter than the prior CUPUM volumes, the result of a joint decision by the editors and Springer to produce a more focused book. The overall CUPUM community remains vibrant, and many of the submissions that were not included here will be presented as conference papers at CUPUM.

1.1 Overview of the Book

The topics of the book were largely determined by the authors, drawn from the remarkably diverse CUPUM community. Although the authors work in a wide variety of research areas, we found that their submissions fell into three general categories, which we have used to organize the book: Digital Cities, Mobility Futures, and Fine-scale Urban Analysis. Next we provide a brief overview of each of these parts.

1.1.1 Digital Cities

How are digital technologies embedded into cities? The first part features four chapters tackling different aspects of this fundamental question. In recent years there has been a continued decline in interest in the topic of "smart cities" in the CUPUM community, and the chapters here reflect a desire to move beyond the earlier idealism of smart city concepts to research that engages with real-world applications. Chapter 2, "Hybrid Smartness: Seeking a Balance Between Top-down and Bottom-up Smart City Approaches," by Mennatullah Hendawy and Iasmin Fernanda Kormann da Silva, provides a useful overview of the many related terms and concepts, suggesting we move beyond discredited top-down smart city concepts and pursue a more nuanced "hybrid smartness." Chapter 3, "Interpreting the Smart City Through Topic Modeling," by Zhibin Zheng and Renée E. Sieber, continues the trend of focusing on empirical examples through the analysis of a large corpus of applications to the Canadian Smart Cities Challenge, revealing a more nuanced, contextual understanding of smart cities than is found in popular definitions. Whereas many urban planners and scholars have focused on the potential for smart cities to realize greater sustainability and efficiency, in practice cities have invested heavily in technologies for surveillance and social control. Chapter 4, “The Venue Code: Digital Surveillance, Spatial (Re)organization, and Infrastructure Power During the COVID Pandemic in China,” by Xialing Chen, provides a gripping ethnographic account of one component of China’s recently ended set of zero-COVID policies, the Venue Code, where digital tools were used to control COVID but at great cost to personal freedom.

Even as the allure of the smart city has faded, the deployment of digital technologies in cities has continued unabated. Therefore, in the place of smart cities, Pamela Robinson and Peter Johnson in Chap. 5, “The Platformization of Public Participation: Considerations for Urban Planners Navigating New Engagement Tools,” introduce the concept of *platforms* as a useful rubric for describing many of these systems, which connect diverse users through a shared digital infrastructure. Focusing on a suite of new digital public participation tools that experienced growing popularity during the pandemic, they call attention to the ways in which these tools may be constraining and narrowing public participation, a fundamental component of urban planning practice.

1.1.2 Mobility Futures

Among the many types of urban platforms that have emerged, one of the most visible has been *mobility platforms*, such as companies offering app-based scooter and bike rentals and taxi rides, now available in virtually every large city worldwide. The book’s second section contains six chapters related to mobility futures, many of them directly related to the emergence of mobility platforms.

The section kicks off with Chap. 6, by Zhenpeng Zhou, “Shared Micro-mobility: A Panacea or a Patch For Our Urban Transport Problems?” This chapter reviews the rise of “micro-mobility” services such as bike and scooter apps, raising tough questions about how they are best incorporated into our urban mobility system. Urban platforms produce data, and in the case of mobility, these data have often been made available to researchers to plumb for new insights. The next two chapters each make use of data from bike-sharing systems to analyze bikeability in two cities. Chapter 7, “Understanding Bikeability: Insight into the Cycling–City Relationship Using Massive Dockless Bike-sharing Records in Beijing,” by Enjia Zhang, Wanting Hsu, Ying Long, and Scott Hawken, mines bike-sharing data to analyze biking in China’s capital. Chapter 8, “Disclosing the Impact of Micro-level Environmental Characteristics on Dockless Bikeshare Trip Volume: A Case Study of Ithaca,” by Qiwei Song, Wenjing Li, Jintai Li, Xinran Wei, and Waishan Qiu, conducts a related analysis for a small U.S. city.

Whereas descriptive research highlights the characteristics of where people bike, Chap. 9, "A Planning Support System for Boosting Bikeability in Seoul," by Madiha Bencekri, Donggyun Ku, Doyun Lee, and Seungjae Lee, extends that question into how cities can be made more bikeable, through a rigorous modeling exercise in Seoul. In addition to bike-sharing, ride-hailing apps that provide users the ability to request a taxi have become ubiquitous, but many questions remain about the costs and benefits of these services. Chapter 10, "Integrating Big Data and a Travel Survey to Understand the Gender Gap in Ride-hailing Usage: Evidence from Chengdu, China," by Si Qiao, Anthony Gar-On Yeh, and Mengzhu Zhang, provides an interesting analysis through a gender lens of who is using ride-hailing services in Chengdu, China.

While micro-mobility, bike-sharing, and ride-hailing are all part of our global mobility reality, in recent years investment has flowed into companies seeking to create *advanced air mobility services*. The general concept combines the aeronautical technologies used by drones with the business models of the ridesharing platforms to create a totally new mode of travel—by small electric-powered aerial vehicles. Although focused on the technical details of urban airspace route planning, Chap. 11, "Urban Airspace Route Planning for Advanced Air Mobility Operations," by Xi Wang, Perry Pei-Ju Yang, Michael Balchanos, and Dimitri Mavris, incidentally demonstrates the regulatory complexity of urban airspace. As a result, urban planners and analysts may have a major role in extending the domain of transport planning to the skies to accommodate this technology into urban transport systems.

1.1.3 Fine-Scale Urban Analysis

The final part, on fine-scale urban analysis, contains submissions that demonstrate the potential for the innovative use of data for urban analysis. Many of these chapters highlight the use of big data, or innovations in established methods such as planning support systems (PSS) and cellular automata. Chapter 12, "'Eyes on the Street': Estimating Natural Surveillance Along Amsterdam's City Streets Using Street-level Imagery," by Timo Van Asten, Vasileios Milias, Alessandro Bozzon, and Achilleas Psyllidis, uses newly available street imagery and computer vision techniques to use data to test longstanding urban theories. Chapter 13, "Automatic Evaluation of Street-level Walkability Based on Computer Vision Techniques and Urban Big Data: A Case Study of Kowloon West, Hong Kong," by Lu Huang, Takuya Oki, Sachio Muto, Hongjik Kim, Yoshiki Ogawa, and Yo-shihide Sekimoto, applies similar techniques to analyze the walkability of the Kowloon West, Hong Kong. Both of these chapters illustrate how the availability of visual data—and the related techniques to process them—hold great potential for the analysis of urban form.

One longstanding area of research in the CUPUM community has been the development and study of planning support systems, which link various data and models to users engaged in planning and decision-making. Chapter 14, "Promoting Sustainable Travel Through a Web-based Tourism Support System," by Yidai Kato and Kayoko

Yamamoto, contains an intriguing example of a novel PSS that puts tourism information at visitors' fingertips to encourage more sustainable tourism patterns. Australia's Urban Research Infrastructure Network (AURIN) encompasses not only a PSS but also rich geospatial data. Chapter 15 features a study involving AURIN, "Applying the AURIN Walkability Index at the Metropolitan and Local Levels by Sex and Age in Australia," by Arsham Bassiri Abyaneh, Andrew Allan, Johannes Pieters, Sekhar Somenahalli, and Ali Soltani. The chapter reports the result of an analysis exploring the relationship between an existing walkability index and walking-related travel behaviors.

The book's final two chapters contain innovative examples of spatial modeling, long a key component of many PSS. Chapter 16, "Predicting Urban Heat Island Mitigation with Random Forest Regression in Belgian Cities," by Mitali Yeshwant Joshi, Daniel G. Aliaga, and Jacques Teller, uses machine learning methods and remote sensing to find that green roofs effectively mitigate urban heat, a timely finding for planners and policymakers facing a hotter future. Chapter 17, "A Framework to Probe the Uncertainties in Urban Cellular Automata Modeling with Multilevel Density for the Wallonia Region, Belgium," by Anasua Chakraborty, Ahmed Mustafa, Hichem Omrani, and Jacques Teller, contributes to the cellular automata tradition of urban growth modeling by analyzing how cell sizes and other decisions affect the uncertainty of predictions.

1.2 Intelligence for Future Cities

When deliberating on an appropriate title for the book, we reflected on broadly shared themes among the chapters. Many of the chapters use novel data and modeling techniques to arrive at new insights about a range of topics of interest to planners, from mobility to urban heat islands. Unlike more purely descriptive "urban science" seeking to identify universal urban laws, most of these chapters reflect various agendas for urban change. The goal of these studies is not only to understand how cities function today but also to point the way toward cities that are more walkable, more bikeable, more participatory, and more sustainable. As a result, we decided on the title "Intelligence for Future Cities: Planning Through Big Data and Urban Analytics." Ironically, at the time of this writing there was great interest in advances in *artificial* intelligence (AI), such as AI models that could generate images from text prompts or provide responses to questions in a chat. Although we have no doubt these tools will in time find their way into our field, in our title we refer to a more traditional concept of intelligence linked to decision-making.

In his classic work *The Public and Its Problems*, the American philosopher John Dewey defined intelligence as the anticipation of the consequences of actions (Dewey 1927). Furthermore, he argued that we should understand intelligence as not simply an individual attribute but as something shared by a community. This definition should resonate with many in the CUPUM community, involved in efforts to conduct analysis that ultimately will be used to test alternative proposals and anticipate the future cities they may create. We observe that most of the chapters in this volume feature multiple authors, and many of the scholars are deeply involved in local infrastructures and initiatives, ensuring their analyses are useful for real-world decisions. In our politically and socially fragmented societies, cities remain a key locus for shared life. Although we should expect no end to disagreements about city goals, the shared work of the CUPUM community is to build data and tools that boost the levels of intelligence with which we manage and plan our cities. This means not only demonstrating new methods and findings but also translating them into shared tools like PSS. It is our hope that this volume, in addition to the CUPUM conference, provides a modest contribution to this goal.

Reference

Dewey John (1927) The public and its problems. Ohio University Press, Athens, p 12

Part I
Digital Cities

Chapter 2
Hybrid Smartness: Seeking a Balance Between Top-Down and Bottom-Up Smart City Approaches

Mennatullah Hendawy and Iasmin Fernanda Kormann da Silva

Abstract The idea of smart cities has been the subject of a range of views in recent years. A review of the terms and concepts around the "smart city" reveals three distinct ways of making urban development "smart": a techno-centric approach (top-down), a social-centric approach (bottom-up), and a socio-technical approach (in-between). While each direction is significant in its own way, we argue that it is crucial to support an alternative, integrated idea of the smart city through "hybrid smartness." In this new approach to smart cities, both top-down and bottom-up smart ideas are balanced, and recognized. It is the interweaving of both ideas that ensures an urban development model that promotes equity and sustainability. This hybrid thinking approach can thus inform the wider planning discourse that is usually trapped between top-down and bottom-up extremes.

Keywords Smart city · Intelligent city · Digital city

2.1 Introduction

Cities are complex urban systems that serve as the foundation for social and economic development and innovation (Batty, 2009). The increasing dynamism and multiplicity of urban phenomena have accelerated the need to seek alternative and more effective ways to deal with the new challenges (Nam and Pardo, 2011). The idea of "smart cities" was introduced into this context as a concept used to describe technological responses to such urban issues (Chourabi et al., 2012; Coletta et al., 2019). Between the early 1990s and 2010, academic interest in smart city research rose gradually, and after 2015, it has expanded significantly (Sharifi et al., 2021). The term has been

M. Hendawy (✉)
Center for Advanced Internet Studies, Center for Advanced Internet Studies GmbH (Germany), Impact Circles e.V. (Germany), Ain Shams University (Egypt), Cairo, Egypt
e-mail: mennatullah.m.hendawy@tu-berlin.de

I. F. K. da Silva
Impact Circles e.V. (Germany), TU Dortmund University (Germany), Dortmund, Germany

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_2

discussed among academics, urban planners, and decision-makers for over three decades as a solution to serious global urban concerns.

Since its emergence, the term "smart city" has been the target of a variety of visions and imaginations of what and how such cities are or should be (Ching and Ferreira, 2015). The proposed definitions seek to establish the qualitative distinctions as well as the technical, infrastructural, and performance characteristics that a smart urban environment should possess. We notice that scholars are progressively shifting their focus from technology and infrastructure to people and communities. Sharifi et al. (2021), in examining the thematic evolution of smart city research since 1991, show that the topic emphasis of the field can be broken down into three major clusters: the concept itself, the Internet of Things, and big data analytics. Additionally, Sharifi et al. (2021) note that from 2015 on, new concerns have received more consideration, such as sustainability and climate change, decision making and governance, security and privacy, and data management and operational optimization. Furthermore, since the introduction of smart cities, we have observed that their idea has shifted from a sectoral to a multifaceted approach, in such a way that places city management at the forefront and emphasizes the engagement of many stakeholder groups in municipal administration.

So far, however, smart city ideas and emerging terms are implemented as a collection of smaller initiatives as opposed to strategic, integrated ways to manage the city and its infrastructure (Attaran et al. 2022). The emerging term "smart city" emerged initially predominantly in the engineering and computer science fields, In this chapter, we suggest that these techno-centric approaches to smartening sites are "top-down" manifestations of smartness. Nevertheless, many scholars have challenged the conventional, techno-centric notion of a smart city by arguing that it promotes an undesirable picture of a future city characterized by centralized computer monitoring and control, which will end up not serving its residents' interests (for example, see Costa and Oliveira, 2017; Green, 2019; Greenfield, 2013). Only recently, alternative visions of the smart city have arisen advocating for greater emphasis on users' needs, desires, and perspectives (for example, see Pype, 2017). In this chapter, we suggest that these emerging (more social) perspectives represent a new "bottom-up" tradition in smart cities discourse.

So a general reading of the development of the term "smart city" shows two different ways of dealing with smartening urban development: techno-centric (top-down ideas) and socio-centric (bottom-up ideas). In between these two approaches, there is a forgotten socio-technical one. From this premise, we would like to extend the spectrum of defining smart cities. Scholars argue that the way things are defined constructs their reality (Berger and Luckmann, 1967). In the smart city discourse, Hollands (2008) argues that the smart city idea becomes a self-imposed label and a branding tool for cities. Moreover, Murgante and Borruso (2014, 748) claim that in smart cities, associated programs are "based on ICT improvement," which can result in a loss of focus on the underlying urban issues.Borruso adds that it is crucial that programs start by identifying the beneficiaries of a particular project. In light of these debates, several scholars call for "alternative" ways of understanding and dealing with the smartening of cities (i.e., Cardullo and Kitchin, 2018; McFarlane and Söderström, 2017, p. 1). For instance, the UN Economic and Social Council report mentions that there is a growing need to localize smart infrastructure (UN, 2016).

Ultimately, we assume that technology must be conceived not as a pragmatic answer in terms of its technical aspects or its ability to solve a particular problem, but as a means to be implemented in tandem with socio-economic, legal, environmental, and political goals and priorities. As such, the process of digitalization involves the pervasive integration of digital communication and media technologies within various domains of social existence, resulting in a reconfiguration of societal norms and practices (Srai and Lorentz, 2019). Our hypothesis is that by taking cues from the "hybrid" integration of different top-down and bottom-up ways of being "smart" in the city, one can harness and merge the potential of techno-centric and socio-centric definitions in order to establish a new approach to digitalization and planning that is more hybrid.

Towards this end, Sect. 2.2 of this chapter elaborates on the idea of the smartening of cities by investigating the terms and concepts of smart cities that have emerged over time. Sections 2.3 and 2.4 categorize and explain the top-down and bottom-up ideas for smartening cities, respectively. Section 2.5 then elaborates on our main argument: the need for hybrid smartness ideas in which both meet to achieve a balanced digital transformation in societies. Following, in Sect. 2.6, we discuss the way forward and the impact of hybrid smartness ideas to planning. Finally, in Sect. 2.7, we draw our conclusions.

2.2 The Idea of Smartening Cities-Exploring Smart Cities' Terms and Concepts Over Time

In order to explore smart city concepts and terms, we searched for the names used to describe smart cities over the years. The search methodology utilized in the current investigation was not systematic in nature. Rather, whenever we encountered a term, it was added in an ad hoc manner to the table. The primary objective of this approach was not to compile an exhaustive list of all possible terms that attempt to define smart cities; instead, the focus was to present a concise account of the relevant terminology that can be associated with the top-down and bottom-up approaches to smartness. Notably, future research could employ a more structured approach to identifying terms in order to produce a more exhaustive list of terms.

Although the term "smart city" emerged in the 1990s, the search revealed that concept's roots go back to the 1960s. We identified 20 terms and concepts linked to the idea of the smart city.[1] In this section, we reflect on those emerging terms and

[1] The terms and concepts searched are not exhaustive and new ones may arise constantly. It is worth noting that, due to the focus of our research, we intentionally excluded the environmental related smart city terms and concepts, which would demand distinct notions and future research.

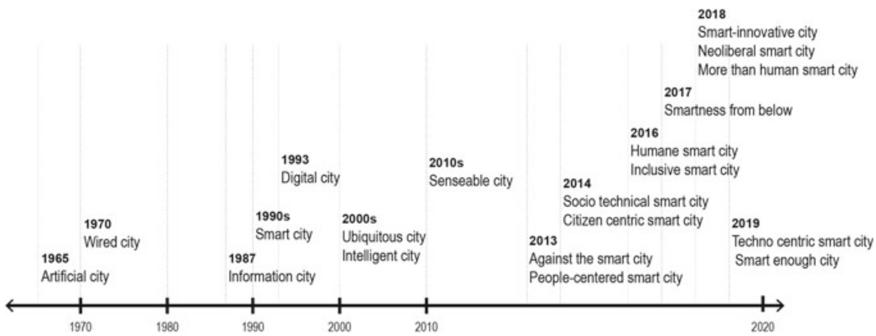


Fig. 2.1 Timeline of the terms and concepts linked to the idea of the smart city (*source* authors)

concepts, and share their definitions. Figure 2.1 shows the emergence of the terms and concepts searched over the years.

The idea of smart cities can be traced back to 1965, when the term "Artificial City" was introduced and coined by Alexander (1965). This term was used by him to describe "cities and parts of cities which have been deliberately created by designers and planners" opposing the idea of a city that arises spontaneously over time (Alexander, 1965, p.152). Later, in the 1970s, the term "Wired City" emerged to denote a vision of society centered on communication and to advocate the use of emerging computer and telecommunications technologies in cities (Dutton, 2019). In 1987, the term "Information City" was used to reinforce the centrality of information and new information and communication technologies to socioeconomic development (Hepworth, 1987).

From the 1990s on, early appearances of the term "Smart City" were used to describe urban contexts with a focus on the significance of new ICT with respect to modern infrastructure within cities (Albino, Berardi, and Dangelico, 2015). At the same time, in 1993, an initiative of hackers and cyberspace activists in Amsterdam outlined the term "Digital City"[2], which was meant to be a metaphor for public participation. Thus, "Digital City" refers to a "digital public space to be used to improve the dissemination of political information and communication between citizens and politicians" (van den Besselaar and Beckers, 2005, p. 68).

In the 2000s, the term "Ubiquitous City" appeared mainly in big project proposals in South Korea. It advocated a vision that incorporates the concept of ubiquitous computing in designing urban infrastructure development (Shin, 2009). In the same period, the term "Intelligent City" emerged with a focus on "creating an environment supporting technology, innovation, learning, and knowledge-based development" (Komninos, 2002). Later, in the 2010s, the term "Senseable City" was coined, and used by the influential MIT Senseable City Lab. It brings a vision of a city where sensors and hand-held electronics allow new approaches to checking on activities

[2] *De Digitale Stad* in the original Dutch language.

or changes that are happening in the city (Martino et al., 2010; Ratti and Claudel, 2016).

In the year 2013, we noticed that there was a shift in the definitions, and, from that point on, several new terms emerged to draw attention to and accentuate different concepts related to Smart Cities. Greenfield (2013) used the title "Against the Smart City" to bring forth a critique of the smart city, one that would truly respond to the will of all citizens and that would incorporate and respond to the interconnected, intricate, and human complexities of the urban environment. In the same year, a group of researchers (Giovannella et al., 2013) used the term "People-Centered Smart City" to also advocate a vision driven by the dimensions of human complexities advocating a bottom-up perspective that puts people in the central position of strategy formulation. In a similar plea, one year later, Lee and Lee (2014) used the name "Citizen-Centric Smart City", stating that "smart city services would be better developed following the actual and precise needs of citizens" (Lee and Lee, 2014, p. 9).

At the same time, Carvalho (2014) used the term "Socio-Technical Smart City" bringing in a more holistic view on smart cities: they are shaped by social, economic, and technological variables. According to Carvalho (2014), "Socio-Technical Smart Cities" consider citizens, corporations, and local governments' needs, expectations, and viewpoints; and should be viewed as a continuous process of experimentation and adaptation, rather than a fixed and predetermined outcome.

Returning to the focus on people, later, de Oliveira Neto and Kofuji (2016) brought to light the idea of an "Inclusive Smart City" to update the traditional image of smartness and include accessibility issues and people with disabilities' needs as part of the smart city agenda. In this sense, Pype (2017) described "Smartness From Below", which consists of approaching smartness "through the ways in which... [local inhabitants] deal with innovation, technology, and creativity and talk about ways of creating knowledge, tools, and practices necessary in urban life" (Pype, 2017, p. 99).

Costa and Oliveira (2017) used the term "Humane Smart City" to advocate that smart cities should be designed to prioritize the well-being and quality of life of citizens, rather than just economic growth and efficiency. Moreover, Thakkar (2018) called for a "People-Centric Smart City" when he sought to analyze the smart initiatives that are part of the Smart City Mission of the Government of India, concluding that "the most vital aspect of the Smart City Mission should be the citizens who live and work in these cities, and they must be integral to the implementation process as well" (Thakkar, 2018, p. 19). Also in 2018, the term "Smart-innovative City" was coined "to emphasize the specific focus on smart city strategies directed at facilitating innovation and driving economic growth in turn" (Maalsen, Burgoyne, and Tomitsch, 2018). At the same time, aiming from a critical point of view, Cardullo and Kitchin (2018) used the term "Neoliberal Smart City", arguing that the smart city discourse is heavily influenced by neoliberal ideology, which stresses economic growth over the interests and concerns of inhabitants. In 2019, also from a critical perspective, de Souza, Hunter, and Yigitcanlar (2019) used the term "Techno-Centric Smart City" to describe the smart cities created by partnerships between local governments and technology companies. During the same time, Green (2019) used the term "Smart

Enough City" to designate that "being 'smart' is a means rather than an end, and that the focus can rightfully turn to the social needs that technology addresses" (Green, 2019, p. 157).

This review of smart-city-related terms is summarized in Table 2.1, which lists the terms and categorizes their approaches. When considering the development of smart city narratives over time, we notice three approaches to smart cities: (a) top-down smartness: those that are techno-centric and implemented in a top-down fashion by those in charge; (b) bottom-up smartness: those that are people-centric and implemented in a bottom-up fashion based on citizen input regarding problem definition and program design, and (c) in-between/hybrid socio-technical smartness: those that advocate for a balance between the top-down and bottom-up ideas. It is important to note that the "in-between smartness" idea proposes a viewpoint that is consistent with our argument of the need for a balanced hybrid idea to smartness that is presented in this chapter. However, as seen in the table, it is not as highlighted in literature as much as the other two approaches.

In the next sections, we will illustrate the top-down and bottom-up ideas for smartening cities and, in turn, the need for a balancing hybrid narrative for smart cities.

2.3 Early Ideas: Top-Down Smartness

In this section, we provide a more detailed description of techno-centric approaches we call top-down smart cities.

The central purpose of the term "smart" in the top-down idea is mostly related to equipping the city with digital infrastructure (sensors, smart grids, AI, etc.). We notice that most of the discussion that falls under this idea emphasizes *how* to make equipment, services, tools, tasks, etc. smarter (i.e., reduce consumption of time, effort, and resources); however, not enough attention is given to *why* to do that. Thus, a gap was created between technical development and its application in people's daily lives.

In this regard, the following ideas about the smart city can be considered top-down smartness: Artificial City (Alexander, 1965), Wired City (1970s), (see Dutton, 2019), Information City (Hepworth, 1987), Smart City (1990s), Ubiquitous City (2000s), Intelligent City (2000s), Senseable City (2010s), Neoliberal Smart City (Cardullo and Kitchin, 2018), Smart Innovative City (Maalsen et al., 2018), and Techno-centric Smart City (de Souza et al., 2019). This tracing shows that the early ideas of smart cities were primarily related to computing and digitalization (whereas the smart city term itself emerged more recently). The initial definitions were focused on analyzing and defining the transition of cities with the use of new information and communication technologies (ICTs). Therefore, the initial concepts aimed to bring the discussion about the new opportunities of these technologies to the urban arena and advocate the narrative that cities should adopt new technologies.

We postulate that the success parameters of smart cities in top-down ideas of smart cities are usually quantitative in nature, which makes it hard to acknowledge

Table 2.1 Tracing smart city-related terms over time since their emergence in 1965 to 2022 (authors) (The reference column means that it the source of the term when a term was first mentioned to the best of our knowledge)

Term	Year	Definition	Reference/source	Approach
Artificial city	1965	"Cities and parts of cities which have been deliberately created by designers and planners"; in opposition of cities which "have arisen more or less spontaneously over many, many years".	Alexander (1965)	Top-down
Wired city	1970s	Vision of a society centered on communication. Use of emerging computer and telecommunications technology to collect data to drive the economic and social development of the city and society.	Dutton (2019)	Top-down
Information city	1987	Vision of information and of new information and communication technologies as central axes of structural changes in the economy and in society.	Hepworth (1987)	Top-down
Smart city	1990s	"The focus was on the significance of new ICT with regard to modern infrastructures within cities."	Albino et al. (2015)	Top-down
Digital city	1993	City metaphor for public participation; it is "a digital public space to be used to improve dissemination of political information and communication between citizens and politicians."	van den Besselaar and Beckers (2005)	Bottom-up
Ubiquitous city	2000s	Incorporates the concept of ubiquitous computing in designing urban infrastructure development.	Shin (2009)	Top-down
Intelligent city	2000s	Focuses on "creating an environment-supporting technology, innovation, learning, and knowledge-based development."	Komninos (2002)	Top-down

(continued)

Table 2.1 (continued)

Term	Year	Definition	Reference/source	Approach
Senseable city	2010s	Vision of a city where sensors and hand-held electronics allows new approaches to check on activities or changes that are happening in the city.	Martino et al. (2010)	Top-down
Against the smart city	2013	It imagines and supports an alternative vision of the smart city, one that would truly respond to the will of all citizens and that would incorporate and respond to the interconnected, intricate, and human complexities of the urban environment.	Greenfield (2013)	Bottom-up
People-centered smart city	2013	A vision "driven by a "person centered in place" design approach supporting the harmonious development of all relevant dimensions of the human experience."	Giovannella et al. (2013)	Bottom-up
Citizen centric smart city	2014	"Smart city services would be better developed following the actual and precise needs of citizens." To execute smart city services across government functions, functions must be cross-referenced and connected.	Lee and Lee (2014)	Bottom-up
Socio technical smart city	2014	Combines people-oriented and technology-oriented approaches.	Carvalho (2014)	In between
Inclusive smart city	2016	Updates the traditional Smart City idea, including accessibility issues and People with Disabilities need as part of the agenda.	de Oliveira Neto and Kofuji (2016)	Bottom-up
Smartness from below	2017	Approach smartness "through the ways in which... [local inhabitants] deal with innovation, technology, and creativity and talk about ways of creating knowledge, tools, and practices necessary in urban life"	Pype (2017)	Bottom-up

(continued)

Table 2.1 (continued)

Term	Year	Definition	Reference/source	Approach
Humane smart city	2017	A city that prioritizes the well-being and quality of life of its residents through social inclusion, equity, participation, and integration, while addressing complex urban development issues and encouraging transparency and citizen participation in governance.	Costa and Oliveira (2017)	Bottom-up
People-centric smart city	2018	"Are about 'collaborative technology' that brings about collaboration among urban communities, citizens and city governments."	Thakkar (2018)	Bottom-up
Smart-innovative city	2018	Emphasizes the "specific focus on smart city strategies directed at facilitating innovation and driving economic growth in turn."	Maalsen et al. (2018)	Top-down
Neoliberal smart city	2018	"Whilst setting appropriate goals for cities via systems of urban benchmarking, the neoliberal smart city aims to attract foreign direct investment, offering areas of the city as testbeds to pilot new technologies, fostering innovative indigenous start-up sectors or digital hubs, and attracting mobile creative elites."	Cardullo and Kitchin (2018)	Top-down
Techno-centric smart city	2019	Partnerships between local governments and technology companies. "Focused on the transformation of the existing technical and physical infrastructures of a city or greenfield urban projects."	de Souza et al. (2019)	Top-down
Smart enough city	2019	"In the smart enough city, where being "smart" is a means rather than an end, the focus can rightfully turn to the social needs that technology addresses."	Green (2019)	Bottom-up

* The first paper mentioning the term could not be found. The added reference is one that mentions the emergence of the term.

the perceptions of the public and the concrete improvement in citizens' quality of life. Even though the definitions of top-down ideas sometimes show concerns about social issues such as the "improvement of the quality of life of citizens"). However, the definitions often lack a clear proposal regarding how to implement this vision.

As such, top-down smartness involves early conceptions of smart cities that often come from governments and private technology companies.or example, the initiative "Smarter Planet" launched by IBM in 2008, promoted the use of advanced data and technologies to improve infrastructure and operations for cities and organizations to become more efficient, sustainable and responsive (see Dirks et al., 2010). In these cases, smart cities are built on the visions and interests of specific actors; they may or may not have a social concern, but in any case, they tend not to integrate other stakeholders in the process (Hollands, 2014). The strategies presented in this perspective are usually led by the state or corporations and later engage the population. In general, they carry out a neoliberal approach with an uneven power relationship concentrated on political and economic actors, who are the ones who dominate decision-making and determine the application and operation of technological infrastructures.

Goodspeed (2014, p. 11) argues that "technology has always had a large role to play in utopian visions of the future," linking the top-down tradition to older concepts of cybernetic systems theory based ultimately on a logic of efficiency and system control. Moreover, Murgante and Borruso (2014) argue that technology can be a valuable tool in making cities more efficient and well-managed, but it should be used as a means to an end, not the ultimate goal. Seen in these terms, top-down smartness lacks strategies to absorb citizens as actual parts of the concept and processes. This might result in the neglect of local particulars and a smart city that is separated and dislocated from the local scale.

2.4 The Emergence of Bottom-Up Smartness

This section describes the more socially-oriented approaches to the smart city that began to emerge around 2013, which we characterize as "bottom-up smartness."

We postulate that ideas of bottom-up smartness shift the central object of concern from (smart) technology to people, supporting a human-centered design of technological systems. As such, the bottom-up smart city terms and concepts seek to emphasize the people using the technology and address societal issues, and the term "smart" becomes mostly related to civic engagement and participation. In this perspective, civic innovation and multilevel collaboration gain traction as ways to address the complex challenges facing smart urban environments.

On this basis, the following ideas about the smart city can be considered as bottom-up smartness: the digital city (1993) , against the smart city (Greenfield, 2013), people-centered smart city (Giovanella et al., 2013), citizen-centric smart city (Lee and Lee, 2014), inclusive smart city (de Oliveira Neto and Kofuji, 2016), smartness from below (Pype, 2017), humane smart city (Costa and Oliveira, 2017), people-centric smart city (Thakkar, 2018), and smart enough city (Green, 2019). It is acknowledged that the ideas of smart city framed here as bottom-up smartness, encompass a diverse array of perspectives, including but not limited to the provision of digital spaces for communication and participation, the facilitation of smart services, the promotion of inclusivity for all citizens, and the implementation of collaborative technology. Our tracing demonstrates that starting in 2013, there began to be a shift in the way smart cities became addressed in a more critical way, starting with the work of Adam Greenfield, titled "Against the Smart City". Greenfield (2013) objected to the dominant technical vision of smart cities because, in his view, such a concept promotes an undesired urban image that includes centralized computational surveillance and control, which the earlier experience of modernism had demonstrated produced unpleasant and dysfunctional cities. He claimed that cities in this sense would not fulfill the interests of their people, giving an example of Masdar City in UAE (Greenfield, 2013).

On the one hand, bottom-up ideas to smartness can be very effective at addressing specific local needs and concerns. On the other hand, bottom-up smartness requires a significant amount of community involvement and engagement, as well as political will and technical compromise, all of which may be challenging to implement in reality. On top of that, not all communities and neighborhoods are equally qualified to lead the development of smart city initiatives (Vanolo, 2013). Consequently, there is a possibility that a bottom-up approach may result in an unequal allocation of resources and services, favoring locations and groups that are better able to drive the development of smart city projects. Bottom-up initiatives in general encompass fragmented and individual projects without wider coordination (Cugurullo, 2017). This may complicate the implementation of large-scale projects and infrastructure and lead to a fragmented and disconnected smart city system overall.

2.5 Towards Hybrid Smartness: The Need for a Socio-Technical Approach

The above tracing of the smart city concepts and terms shows that there are techno-centric (top-down) and socio-centric (bottom-up) approaches to the conception of digitalization in urban development. While each direction is significant in its own way, it is important to note that they also possess specific limitations and problems. Therefore, to achieve a more effective and just implementation of smart city, it is crucial to support an alternative, more integrated idea of the smart city—one that responds to all of its citizens' needs, demands, and desires while at the same time understanding and working with the complex, interconnected, imperfect, and very human realities of urban life. In this section, we focus on highlighting the few ideas for making cities smarter that show a more socio-technical approach that reflects what we frame as hybrid smartness.

Previous smart city projects show that in many cases, technology companies, in collaboration with governmental bodies, construct top-down, techno-centric projects aimed at overall urban service and urban management efficiency. These projects are top-down, big-budget efforts promoted by tech companies that claim improved efficiency and cost-effectiveness in urban service delivery and urban management (Greenfield, 2013). On the other hand, those concerned about 'the people' don't build scalable systems but focus on resolving local concerns using whatever technology and data are available and affordable. They can be implemented in an unstructured and experimental way (Zhou et al., 2021). In this regard, many projects have lower-budget efforts driven by community groups and nonprofits focused on addressing citizen-perceived problems using data and information technology to make their case (see Pype, 2017).

We postulate that there is a need for hybrid smartness approaches that aim at bringing a balance between the techno-centric (top-down) and socio-centric (bottom-up) approaches. In general, smart cities are understood as urban environments capable of satisfactorily executing their economic, social, managerial, and environmental activities due to the smart conciliation between autonomous and conscious citizens (Giffinger et al., 2007). However, as outlined in the previous two sections, the "smart" in smart cities tends to follow either a top-down or a bottom-up approach, neither of which is enough to effectively manage economic, social, administrative, and environmental functions. Here we argue that the optimal approach to smartness would be somewhere between the techno-centric and social-centric approaches.

One scholar who has promoted moving this direction is Carvalho (2014), who proposes an integrated, effective, and dynamic application and development of technology within a city. His research "suggested that more than 'technologising the way out' of current urban problems, a number of socio-technical processes have to unfold [...] so that new IT solutions can effectively challenge current regimes of urban provisions" (Carvalho, 2014, p. 15). However, although this socio-technical approach emerged in 2014, we do not find supporting terms for it in our tracing of the smart city terms (presented in Table 2.1). Figure 2.2 presents an updated timeline showing the shift towards more socially-oriented smart city terms and ideas.

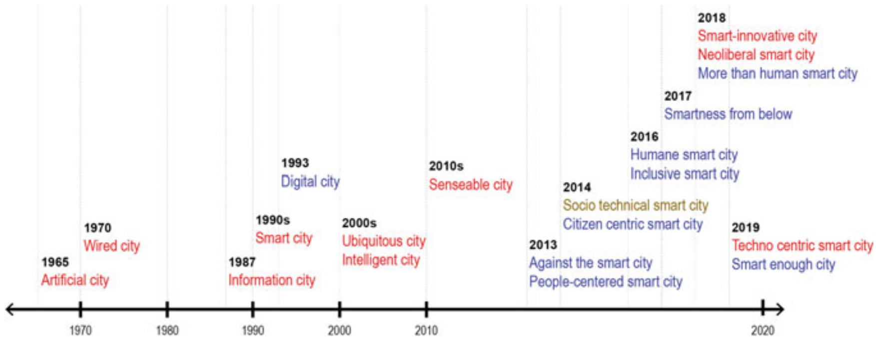


Fig. 2.2 Timeline of Smart city-related terms over time since their emergence in 1965 to 2022; the terms are colored according to their approaches (red=top-down, blue=bottom-up, and yellow=in-between)

In this framework, we propose that the path to a more balanced and equitable urban transformation is to develop a hybrid smartness of smart cities. To achieve this, hybrid smartness may serve as a means of assuring the egalitarian and moral nature of socio-technical systems in urban life. This approach to smart cities recognizes both low-tech and high-tech as smart resources for co-constructing space; therefore, it entails the interweaving of top-down and bottom-up socio-technical solutions that promote fairness for all elements (human, material, and environmental) in and within urbanity. Based on this, hybrid smartness is able to reposition technology as active actors that need to develop side-by-side with socio-economic, legal, environmental, and political positions in order to create equitable urbanism. Therefore, hybrid smartness can help rethink the role of technology in smart cities, shifting the paradigm to balanced socio-technical approaches. To illustrate our argument, we can draw upon the proposition presented by Walravens (2016) in his work, which also endorses the integration of both top-down and bottom-up approaches to create more effective smart cities. The author highlights the living lab approach, which emphasizes scale, sustainability, and stakeholder involvement. In particular, the "City of Things Lab" in Antwerp is a successful case in which a collaborative approach was adopted, involving the public sector, private sector, citizens, and academia (Walravens, 2016).

As argued by Goodspeed (2014, p. 11), "although IT artefacts can play a role in social change, they do not eliminate the social and political dimensions of cities". Therefore, in the hybrid smartness, digital urban development must develop and apply coordinated systems, standards, and protocols at the same time as tech designers and urban planners take into account the social demands and usage of technology systems by users. Such an articulation demands the understanding that top-down and bottom-up smartness are interdependent: the success of one idea or approach lies in the counterpart of the other.

In addition, we encourage the development of hybrid smartness as a means for conventional cities with limited finances to develop in a smart and efficient manner in order to achieve sustainable, capable, and just growth. The concept of a "smart city" is portrayed as the deployment of high-tech IT solutions that burden city budgets, and the benefits of the new real-time data collection are frequently negligible, rendering the smart city concept unsuitable for developing future cities and enhancing the quality of life for their residents.

Furthermore, we assume that smart cities based on hybrid smartness entirely reflect the semantic scope of the word smart: combining different forms of learning in a thorough and extensive process, with feedback and self-improvement in the various arenas. This is a complex yet incremental way to achieve smartness, with milestones, revisions, adjustments, and even setbacks; however, hybrid smartness represents a developmental step towards the desired state of affairs.

We assume that there is an urgent need to change from smart city research and practice to smart and sustainable cities, moving away from purely techno-centric or purely socio-centric approaches in future city design. Through the development of the idea of hybrid smartness and its implementation in a variety of different situations, there is the possibility for a comprehensive, multi-dimensional method for the expansion of cities, as well as for urbanism that is both balanced and egalitarian.

However, we must say that the concept of a hybrid smart city is still a bit undefined and elusive, such an approach would entail the cooperation of multilevel transdisciplinary actors and interdisciplinary experts, providing evidence that the smart city is shaped through collaboration and coproduction. We use this chapter to open up the discussions about these aspects of hybrid smartness, and in the next section, we make an initial attempt to reflect on some cases from reality.

2.6 Reflections on Current Smart Realities

The previous understanding of the different ways of smartening cities (top-down, bottom-up, and in between hybrid approaches) structure the nature of problems addressed and the angle from which digitalisation in cities takes place. In this section, we share some reflections on a few use cases to sort out the differences resulting from top-down, bottom-up, and hybrid ideas.

First, we can take the case of Masdar smart city in the United Arab Emirates as illustrating a big top-down implementation. The project consists of the construction of a city from scratch based on clean energy, compactness, and sustainability. However, the city's socioeconomic effects have been neglected in the debate about the smart city (Kostreva, 2014). Some of the outcomes that may result from this approach were argued to include exclusivity, with only certain members of society able to live in the city, and the widening of socioeconomic gaps (Greenfield, 2020; Kostreva, 2014). Moreover, the implementation of the smart city has already taken enormous economic efforts and is not yet finished (Flint, 2020). This case demonstrates how large-scale strategies, despite their rhetoric of efficiency and cost-effectiveness in urban environments, have a real-world track record of mixed success at best.

Now, examining the city of Kinshasa in the Democratic Republic of the Congo as a use case for bottom-up smartness. Pype (2017) illustrates how the concept of "smartness" depends on who uses it and how it is connected to different objects and imagined users. In this case, smartness can have different meanings depending on the context in which it is used and the person using it. In this sense, she discusses the various connotations and meanings of terms such as "smartness," "intelligence," "knowledge," "creativity," and "innovation" in the context of technological and discursive practices on the local scale (Kinshasa). She examines how these terms are used and understood in different social circles and how they relate to the local and global context. Ultimately, she concludes that understanding the local approaches to knowledge acquisition and information ownership in African cities is essential to the study of technology transfer in Africa. Based on ethnographic work in Kinshasa between 2003 and 2017, Pype (2017) attempted "to sketch the sociohistorical contours of advances in technology and various forms of engagement with tools, scientific knowledge, and technological expertise in an African city" (p. 97). One of the examples in the case of Kinshasa, is the installation of traffic robots which were appreciated by the society not for their efficiency in managing traffic but rather because they replaced policemen who could intimidate or corrupt drivers. As such, the case of Kinshasa exemplifies the significance of small-scale methods and local understanding—which results from a bottom-up idea—when implementing smart city projects.

Finally, to explain more about the hybridity of smart cities, we share some hypothetical examples. A first example is garbage collection and snow plowing; the management of both invites a system-design approach aimed at efficiency. In trying to minimize the cost of waste collection, one risks ignoring the effect of waste collection strategies on what waste is generated, whether it 'should' be separated for recycling, and whether the level of service differs widely across communities. Therefore, a socio-centric approach (bottom-up) that builds buy-in from citizens might be 'better' but it may well still require a large, complex system (techno-centric, top-down) to work effectively. Another example that we can consider to understand hybrid smartness is land use permitting. An efficient digital system might instantly determine whether a proposed development meets all zoning requirements. However, it is critical that there be time available for public comment and review before issuing permits for major or new land use changes. Furthermore, with the availability of an adequate digital infrastructure, the requirement for a more understandable or faster environmental impact analysis can be implemented. In general, the movements towards hybrid socio-technical smartness can be traced back to the 1990s (although these ideas did not gain the same traction as top-down and bottom-up ideas as seen in Table 2.1 and the above analysis) with 'public participation GIS' efforts (PPGIS). It used lower-cost desktop GIS software and newly available digital datasets to analyze and question the environmental impact or community benefits of urban development plans. Ultimately, we can discuss current citizen participation platforms and the development of socio-technical systems that are backed up by technical standards and political support as one of the steps towards hybrid smart cities. In this context, the design of participatory processes based on digital platforms must be able to engage citizens to overcome digital exclusion. Thus, a hybrid approach is crucial to establishing useful instruments for fostering civic engagement and promoting government accountability.

2.7 Final Remarks—Rethinking Planning at Large

Our perspective about the best approach for smart cities is the one that merges and balances both top-down and bottom-up ideas, therefore a hybrid-smartness approach. We started this analysis ourselves, coming from the field of urban planning. Now we would like to reflect back on the planning discourse.

In this section, we share our view as urban planners on how this hybrid thinking and understanding of cities is needed not only in digital but also in the planning of cities at large. In this regard, we argue that speaking about hybridity is needed in planning at large not only in smart cities

There is already a countercurrent debate between top-down and bottom-up approaches that is ongoing—for example Zhou et al. (2021) discuss innovative planning mechanisms implemented at the neighborhood level in China, which supports the integration of top-down and bottom-up initiatives. The authors suggest that an ideal approach for smart city development is to promote an open-source top-down system that incorporates bottom-up applications. We expand these debates by discussing smart cities and the digitalisation of urban environments. The hybrid smartness approach (and the hybrid thinking in general) can be seen as an approach to extending the countercurrent binary approaches of planning; viewing it either from top-down or bottom-up, which could be debunked by introducing the factor of "smartness/digital" to aim for a more balanced understanding of planning and of smart cities. As we notice that the very techno-centric view of the city (top-down) is now followed by a very people-centered view (bottom-up), a balanced view is urgently needed so that "digital people" form and shape cities on an equal footing with analog communities. Hence, hybrid smartness is probably an effective planning-related approach in this context.

2.8 Conclusion

The chapter began with a chronological review of the concepts and terms used in the literature during the past half-century. We comment on the types of 'smartness' approaches envisioned in three cases: technocentric top-down cases, socio-technical approaches, and suggesting a need for a balanced hybrid smartness approach that could capture the beneficial elements of top-down and bottom-up ideas as well as socio-technical approaches.

On the one hand, there are top-down definitions of smartness from the early ideas of concepting smart cities. They provide a narrative coming from governments and tech companies. These conceptions often employ a simplistic view by relying solely on technology to manage urban challenges. Nevertheless, they can present efficient solutions in terms of infrastructure, performance, operation, and upgrading. On the other hand, there are bottom-up definitions that encompass the human dimension in smart cities, which emergence can be traced back to the 'against the smart city' pamphlet by Adam Greenfield in 2013. These conceptions are effective in accounting for societal requirements as well as the ways people (users) interact with technological systems.

We propose to broaden this spectrum, based on the understanding that bottom-up and top-down ideas for smartness are interconnected. In this regard, we present a vision of hybrid smartness, in which semi-technical and socio-technical solutions for urban development are exemplified and articulated in order to merge technical and social approaches to smart urban development. In this new hybrid conception of smart cities, both top-down and bottom-up smart resources are recognized and valued. It is the interweaving of the different ideas that ensures an urban development model that promotes equity and sustainability. We also argue that the hybrid approach allows us to rethink the role of technology within a city and the way planning at large needs to be thought about.

Acknowledgments This article is developed as a part of the research program at Impact Circles e.V. in Berlin, Germany. The idea of focusing on 'hybrid' issues in smart cities builds on the discussions that took place at the Cairo Urban AI project initiated in 2020, in which the first author is one of the founders.

References

Albino V, Berardi U, Dangelico RM (2015) Smart cities: definitions, dimensions, performance, and initiatives. J Urban Technol 22(1):3–21. https://doi.org/10.1080/10630732.2014.942092

Alexander C (1965) A city is not a tree. Arch Forum 122(1):58–62

Attaran H, Kheibari N, Bahrepour D (2022) Toward integrated smart city: a new model for implementation and design challenges. GeoJournal 87(Suppl 4):511–526. https://doi.org/10.1007/s10708-021-10560-w

Batty M (2009) Cities as complex systems: scaling, interaction, networks, dynamics and urban morphologies. Encycl Complex Syst Sci 1041–1071. https://doi.org/10.1007/978-0-387-30440-3_69

Berger LP, Luckmann T (1967) The Social construction of reality. Anchor Books, New York

van den Besselaar P, Beckers D (2005) The life and death of the great amsterdam digital city. Digital Cities III. Information Technologies for Social Capital: Cross-cultural Perspectives. pp 66–96. https://doi.org/10.1007/11407546_4.

Cardullo P, Kitchin R (2018) Smart urbanism and smart citizenship: the neoliberal logic of 'citizen-focused' smart cities in Europe. Environ Plan C: Polit Space 37(5):813–830. https://doi.org/10.1177/0263774x18806508

Carvalho L (2014) Smart cities from scratch? A socio-technical perspective. Camb J RegNs, Econ, Soc 8(1):43–60. https://doi.org/10.1093/cjres/rsu010

Castells M (1989) La ciudad informacional: tecnologías de la información, reestructuración económica y el proceso urbano–regional. Alianza Editorial, Madrid

Ching TY, Ferreira J (2015) Smart cities: concepts, perceptions and lessons for planners. In: Geertman S, Ferreira Jr J, Goodspeed R, Stillwell J (eds) Planning Support Systems and Smart Cities. Lecture Notes in Geoinformation and Cartography. Springer, Cham. https://doi.org/10.1007/978-3-319-18368-8_8

Chourabi H, Nam T, Walker S, Gil-Garcia JR, Mellouli S, Nahon K, Pardo TA, Scholl HJ (2012) Understanding smart cities: an integrative framework. In: 2012 45th Hawaii International Conference on System Sciences. https://doi.org/10.1109/hicss.2012.615.

Clarke R, Heitlinger S, Foth M, DiSalvo C, Light A, Forlano L (2018) More-than-human urban futures. In: Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial - Volume 2. https://doi.org/10.1145/3210604.3210641.

Coletta C, Evans L, Heaphy L, Kitchin R (eds) (2019). Routledge, London, New York

Costa and Oliveira, 2017 Costa EM, Oliveira ÁD (2017) Humane smart cities. The Oxford Handbook of Interdisciplinarity. pp 228–240. https://doi.org/10.1093/oxfordhb/9780198733522.013.19

Cugurullo F (2017) Exposing smart cities and eco-cities: frankenstein urbanism and the sustainability challenges of the experimental city. Environ Plan A: Econ Space 50(1):73–92. https://doi.org/10.1177/0308518x17738535

de Oliveira Neto JS, Kofuji ST (2016) Inclusive smart city: an exploratory study. Lect Notes Comput Sci 9738:456–465. https://doi.org/10.1007/978-3-319-40244-4_44

de Souza K, Hunter M, Yigitcanlar T (2019) Under the hood: a look at techno-centric smart city development. Public Manag 101(11):30–35

Dirks S, Gurdgiev C, Keeling M (2010) Smarter cities for smarter growth: how cities can optimize their systems for the talent-based economy. IBM Institute for business Value.

Dutton WH (2019) Wired city. The wiley blackwell encyclopedia of urban and regional studies. pp 1–4. https://doi.org/10.1002/9781118568446.eurs0414
Flint A (2020) What abu dhabi's city of the future looks like now. Bloomberg. https://www.bloomberg.com/news/articles/2020-02-14/the-reality-of-abu-dhabi-s-unfinished-utopia. Accessed 20 Jan 2023.
Giffinger R, Fertner C, Kramar H, Meijers E (2007) City-ranking of European medium-sized cities. 51st IFHP World Congress "Futures of Cities". pp 1−12.
Giovannella C, Gobbi A, Zhang BX, Perez-Sanagustin M, Elsner J, Fatto VD, Avouris N, Zualkernan I (2013) Villard-de-Lans: a case study for participatory people-centered smart city learning design. In: 2013 IEEE 13th International Conference on Advanced Learning Technologies. https://doi.org/10.1109/icalt.2013.143.
Goodspeed R (2014) Smart cities: moving beyond urban cybernetics to tackle wicked problems. Camb J RegNs, Econ Soc 8(1):79–92. https://doi.org/10.1093/cjres/rsu013
Green B (2019) The smart enough city: putting technology in its place to reclaim our urban future. MIT Press
Greenfield A (2013) Against the smart city: a pamphlet. This is Part I of The City is Here to Use. Do projects.
Hepworth ME (1987) The information city. Cities 4(3):253–262. https://doi.org/10.1016/0264-2751(87)90033-3
Hollands RG (2008) Will the real smart city please stand up? City 12(3):303–320. https://doi.org/10.1080/13604810802479126
Hollands RG (2014) Critical interventions into the corporate smart city. Camb J RegNs, Econ Soc 8(1):61–77. https://doi.org/10.1093/cjres/rsu011
Kitchin, R (2014) Komninos N (2002) Intelligent cities: innovation knowledge systems and digital spaces. Routledge
Kitchin, R (2014) Kitchin R (2014) Making sense of smart cities: addressing present shortcomings. Camb J RegNs, Econ Soc 8(1): 131–136. https://doi.org/10.1093/cjres/rsu027.
Kostreva O (2014) Masdar city focuses on sustainability but excludes the poor. Borgen Magazine. https://www.borgenmagazine.com/masdar-city-focuses-sustainability-excludes-poor/. Accessed 20 Jan 2023.
Lee J, Lee H (2014) Developing and validating a citizen-centric typology for smart city services. GovMent Inf Q 31:S93–S105. https://doi.org/10.1016/j.giq.2014.01.010
Maalsen S, Burgoyne S, Tomitsch M (2018) Smart-innovative cities and the innovation economy: a qualitative analysis of local approaches to delivering smart urbanism in Australia. J Des, Bus & Soc 4(1):63–82. https://doi.org/10.1386/dbs.4.1.63_1
Martino M, Britter R, Zacharias C, Bidermann A (2010) Senseable city: digital urban and modelling. Boston/MA.
McFarlane C, Söderström O (2017) On alternative smart cities: from a technology-intensive to a knowledge-intensive smart urbanism. City 21(3–4):312–328. https://doi.org/10.1080/13604813.2017.1327166
Mitchel WJ (1996) City of bits: space, place and the infobahn. MIT Press
Murgante B, Borruso G (2014) Smart city or smurfs city. In: Computational Science and Its Applications–ICCSA 2014. pp 738–749. https://doi.org/10.1007/978-3-319-09129-7_53.
Nam T, Pardo TA (2011) Conceptualizing smart city with dimensions of technology, people, and institutions. In: Proceedings of the 12th Annual International Digital Government Research Conference on Digital Government Innovation in Challenging Times. https://doi.org/10.1145/2037556.2037602.
Pype K (2017) Smartness from below: variations on technology and creativity in contemporary kinshasa. What Do Science, Technology, and Innovation Mean from Africa? https://doi.org/10.7551/mitpress/10769.003.0008.
Ratti C, Claudel M (2016) The city of tomorrow: Sensors, networks, hackers, and the future of urban life. Yale University Press

Sennett R (2012) No one likes a city that's too smart. The Guardian. http://www.guardian.co.uk/commentisfree/2012/dec/04/smart-city-rio-songdo-masdar. Accessed 10 Oct 2022.
Sharifi A, Allam Z, Feizizadeh B, Ghamari H (2021) Three decades of research on smart cities: mapping knowledge structure and trends. Sustainability 13(13):7140. https://doi.org/10.3390/su13137140
Shin DH (2009) Ubiquitous city: urban technologies, urban infrastructure and urban informatics. J Inf Sci 35(5):515–526. https://doi.org/10.1177/0165551509100832
Shwayri ST (2013) A model Korean ubiquitous eco-city? the politics of making songdo. J Urban Technol 20(1):39–55. https://doi.org/10.1080/10630732.2012.735409
Srai JS Lorentz H (2019) "Developing design principles for the digitalisation of purchasing and supply management," J Purch Supply Manag 25(1): 78–98. https://doi.org/10.1016/j.pursup.2018.07.001.
Thakkar S (2018) Understanding the importance of people centric smart city initiatives. In: UGC Sponsored 14th National Conference on Smart Cities in India: The Road Ahead.
UN-Economic and Social Council (2016) Report of the economic and social council.
Vanolo A (2013) Smartmentality: the smart city as disciplinary strategy. Urban Stud 51(5):883–898. https://doi.org/10.1177/0042098013494427
Varghese P (2016) Exploring other concepts of smart-cities within the urbanising indian context. Procedia Technol 24:1858–1867. https://doi.org/10.1016/j.protcy.2016.05.238
Walravens N (2016) Operationalising the concept of the smart city as a local innovation platform: the city of things lab in Antwerp. Springer International Publishing eBooks, Belgium, pp 128–136. https://doi.org/10.1007/978-3-319-39595-1_13.
Zhou S, Fu H, Tao S, Han Y, Mao M (2021) Bridging the top-down and bottom-up approaches to smart urbanization? a reflection on beijing's shuangjing international sustainable development community pilot. Int J Urban Sci 27(sup1):101–123. https://doi.org/10.1080/12265934.2021.2014939

Chapter 3
Interpreting the Smart City Through Topic Modeling

Zhibin Zheng and Renée E. Sieber

Abstract The Canadian Smart Cities Challenge grant program provided a unique opportunity to investigate what communities across Canada mean when they propose becoming a smart city. We investigated the utility of human-centered machine learning to analyze 137 grant proposals submitted to this program, containing approximately 1.5 million words. We explored whether results generated by topic modeling aligned with or differed from standard smart city definitions in the literature and current urban applications of topic modeling. The analysis resulted in three main findings. First, the prevalence of topics describing rural, regional and Indigenous communities challenged conventional definitions of "city" in the smart city. Second, context (e.g., local culture and language) inferred what constitutes smart, although smartness was not focused on technical innovations. Finally, our abductive approach generated new insights missing from conventional smart city research methods, including 40 percent of finalists and all four winning cities being identified as most representative users of topics.

Keywords Smart cities · Smart cities challenge · Topic modeling · Rural · Indigenous

3.1 Introduction

The smart city remains prevalent in discourses about global urbanization, where data-driven and technology-based solutions promise efficiency, social equality, public safety, greater inclusion and innovation and sustainability (e.g., Albino et al. 2015; Neirotti et al. 2014; Tura and Ojanen 2022). Fueled by these claims, the Canadian federal government issued the Smart Cities Challenge (SCC) grant program in

Z. Zheng · R. E. Sieber (✉)
McGill University, Montreal, Canada
e-mail: renee.sieber@mcgill.ca

Z. Zheng
e-mail: zhibin.zheng@mail.mcgill.ca

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_3

November 2017 (Infrastructure Canada 2017). By April 24, 2018, the federal government received 130 individual or group grant applications from 199 communities. On June 1, 2018, the government announced 20 finalist applications. The finalists then had support from the government to develop a final proposal that was submitted by March 5, 2019 (Infrastructure Canada 2018). On May 20, 2019, Canada awarded four grants, from C$5 million to C$50 million to help winners implement smart city initiatives.

With 137 SCC applications, we had an unparalleled opportunity to extend our knowledge of what constitutes 'smart' and 'city' across an entire country. These applications allowed us to move beyond definitions based on highly-cited cities like Singapore, Boston, or Philadelphia (cf., Ching and Ferreira 2015; Shelton et al. 2015). This corpus also allowed us to expand upon more recent studies, which still tended to examine small numbers of large global smart cities such as London and Melbourne (Clement and Crutzen 2022) or Boston, Helsinki, Seoul and Taipei (Nicolas et al. 2021). We were not the only researchers exploring this corpus. Others have conducted research with the SCC applications but used qualitative methods of a subset, for example examining technology-led versus traditional forms of public participation proposed in the applications (Johnson et al. 2020). We had the opportunity to look across Canada and discover what communities "not in the headlines" considered smart and contrast that with urban communities.

We used a human-centered machine learning approach, in which the role of non-automated interpretation of unsupervised machine learning is explicitly acknowledged, to analyze 137 applications containing approximately 1.5 million words. Traditional methods to interpret the smart city, like literature reviews, content analysis and case studies, may not be the best choice when presented with a large amount of qualitative data (Crowston et al. 2012; Nicolas et al. 2021). Topic modeling, a form of unsupervised classification, allows for a bottom-up, empirical analysis of large volumes of data. We choose the term, abduction, referring to a method that combines bottom-up results emerging from the big data yet is subject to human intervention (Zheng and Sieber 2022). We are guided by three research questions: 1. What happens when the government expands the definition of a city or a smart city, for example related to innovation? 2. Abductively, what constitutes not just a smart city but a smarter city? 3. What is the utility of unsupervised classification for abductive inquiry about smart cities? Do the results align with or differ from the smart city discourse? After a literature review, we describe the implementation details, present results, and discuss findings from this empirical dataset.

3.2 Literature Review

The concept of the smart city is approximately 40 years old (Batty 2013). We should now know what constitutes a smart city; however, the literature has not always clarified the concept (cf., Albino et al. 2015; Clement and Crutzen 2021). "Smart", which early on referred to planned growth as opposed to sprawl (Batty 2013), increasingly overlaps with normative terms like "connected", "participatory", "sustainable", "innovative", "entrepreneurial" or "knowledge-intensive" (Albino et al. 2015). Giffinger et al. (2007) created a widely-used framing of mid-sized European cities into six categories of smartness: smart economy, smart environment, smart governance, smart living, smart mobility, and smart people. Numerous researchers have used the categories to compare cities and refine definitions of smartness (Caragliu et al. 2011; Appio et al. 2019; Nicolas et al. 2021). Nonetheless, the smart city has continued to function as a *tabula rasa*, a "magic concept" (Pollitt and Hupe 2011) onto which anyone could project their aspirations. The smart city becomes eminently adaptable such that it is difficult to determine when a community qualifies as smart.

Technology, particularly information and communications technology (ICT), plays a significant role in characterizing "smartness". How much technology is necessary? Neirotti et al. (2014) divided smart city approaches into "hard" and "soft" domains, distinguished by the degree of dependence on technology. Hard domains like transportation, construction and environmental monitoring are assumed to be more amenable to technologies that allow services to be automated. Soft domains like culture, education and social inclusion presume limited roles for ICTs and are not necessarily directed to, for example, process real-time data. Cohen (2015) suggested that the smart city has evolved away from its high-tech rhetoric and towards human-centered approaches. However, with new disciplines like urban science and concepts like the digital growth machine (e.g., Rosen and Alvarez-Leon 2022), the smart city remains biased towards a technology-dominant discourse, where technical innovation, like big data or artificial intelligence (AI) may prevail over human development (Javed et al. 2022).

Another assumption is that the smart city is a densely developed large city, in large part because the city is the presumed locus of technical innovation. This urban-centered focus is increasingly challenged (Shearmur et al. 2020; Spicer et al. 2021). Urban areas need not be the sole beneficiaries of innovation; rural areas can benefit from technology in many ways including connections to urban centers, improvements of service delivery and increased opportunities for rural residents (Spicer et al. 2021). Shearmur et al. (2020) reminded us of the technological sophistication in rural activities like agriculture. Despite benefits, significant barriers hinder rural areas' goals of becoming smart, including a lack of resources like attracting a skilled workforce or connecting to broadband service (Spicer, et al. 2021). Might a rural area need to emulate the characteristics of the city to be smart? Shearmur et al. (2020, p 311) revealed the conundrum (translated from the French), "[t]he intelligent rural area therefore begins to be considered, even if, paradoxically, it forms part of a contest that promotes the smart city. That is, the rural world will become intelligent only if it urbanizes."

Smart city equals urban reveals an undercurrent of universalism in definitions. That is, a smart city connects technology with the interests and the well-being of communities that is context-independent (Dameri 2013). Many researchers have acknowledged the unique cultural identity of cities that must be accounted for in smart city implementation (Han and Hawken 2018). Obviously, Barcelona, an often-cited example, is not Columbus, Ohio but that may not prevent technology firms from presenting one-size fits-all technological solutions. Part of the uniformity is the creation of standards so cities can be compared in terms of how smart (innovative, creative, open) they are. Cities with varied identities pursue smartness for different reasons. For example, smart cities in Asia, Europe or North America look to share data, increase innovation, improve e-services or adopt green policies. In contrast, smart cities in Central and South America and Africa look to attract foreign investment, move to a knowledge economy, or enhance ICT access in rural areas (Cocchia 2014). Smart cities exhibit rich and diverse practices rather than a single-dimensional race to the top of some world city rankings (Han and Hawken 2018). The tension between universality and particularity is one aspect that we hope can be revealed through our method.

The smart city therefore is an ambiguous concept that complicates standardized analysis (Mosannenzadeh and Vettorato 2014). We neither know precisely what constitutes smart, or smarter, nor do we know what constitutes a city. Analysis of the SCC proposals provides us with an opportunity to empirically assess what we can learn about cities and smartness from 137 Canadian communities.

3.3 Application of Human-Centered Topic Modeling

To understand how communities in Canada conceptualize smart cities, and by implication, how the Canadian federal government envisions the smart Canadian city, we employ human-centered topic modeling (Zheng and Sieber 2022). Figure 3.1 shows our methodology. Topic modeling is a well-known technique for mining text (Jacobi et al. 2016) and for examining urban policy (Debnath and Bardha 2020; Papadopoulos and Charalabidis 2020). Topic modeling is considered heuristic because it learns how the words cluster. Clusters of words become the corpus's themes or topics, although one needs to label and thus assign meaning to the clusters. The problem in topic modeling's application in smart city analysis is that too many authors hide all the judgment calls made to use the method or undervalue the iterative nature of these interventions.

Clement and Crutzen (2021) conducted a similar method as ours (i.e., topic modeling with the Gensim library of long-form policy documents) on two cities to argue that there is no universal smart city policy or definition. Constraining their models to five topics each, they found that cities instead focus on smart city-type topics that align with local policies. Nicolas et al. (2021) also used topic modeling on long policy documents, this time on four cities. They fit resulting clusters into Giffinger et al.'s (2007) six categories. They suggested that categories could be further reduced to a binary: hard and soft disciplines (Neirotti et al. 2014; Albino et al. 2015).

Gan and Qi (2021) remind us that "if the number of topics selected is too small, the meaning under each topic will be too broad." The above approaches can 'shoehorn' what essentially are clusters of clusters into small numbers of categories. Nuance

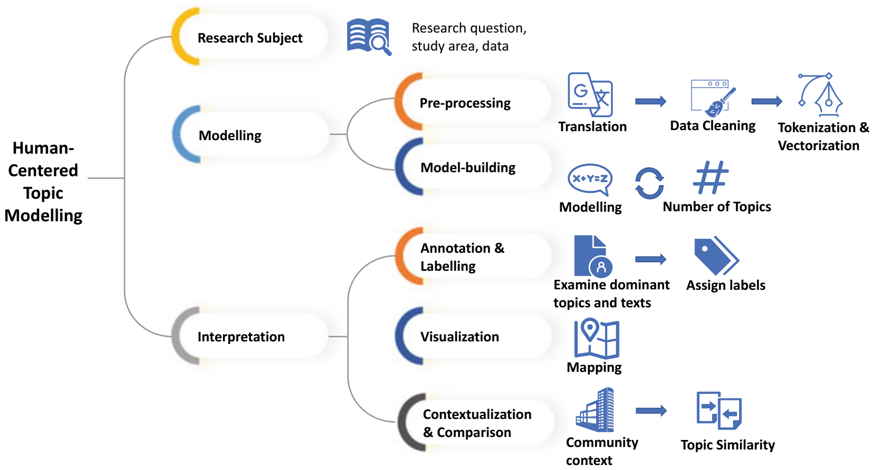


Fig. 3.1 Human-centered Topic Modeling (modified from Zheng and Sieber 2022)

is the key advantage of unsupervised classification. Limiting categories treats topic modeling, a form of unsupervised classification, like supervised classification, in which one classifies their corpus based on a pre-selected set of topics and human-annotated samples that have trained the model.

A human-centered model requires we reflect on the challenges of labeling that captures local context (i.e., Canada). We argue that topic modeling requires iteration between the automation and the (human) interpretation. In some ways, this method resembles traditional qualitative coding methods for examining smart city texts (Johnson et al. 2020). Our method allows us to digest large volumes of text while still being able to qualitatively identify and focus on specific elements in texts. We also introduce a method to conduct cross-corpus comparisons, which have yet to be formalized in urban or smart city research.

Our study analyzed a corpus of 137 proposals, involving nearly 200 communities, 117 of which were SCC grant applications and 20 were final SCC proposals. We did not include 13 additional initial SCC proposals, which were neither available online nor from the applying communities. We generated two topic models, a grant application model (GAM) and a final proposal model (FPM). Since a finalist, who was admitted to the second stage, also submitted a grant application, separate corpora prevented double counting the responses. We operated under the hypothesis that differences between the original grant applications and final proposals could imply what communities viewed as smart and what the federal government viewed as smart. An additional value of these documents was that they existed in a relatively standardized format since communities followed the guidelines from the federal government. The application guide requested four portions: 1. Community information (name, population and prize category), 2. Preliminary proposal (challenge statement, outcomes, activities and projects), 3. Other requirements (proposal summary, commitment, consent) and 4. Survey questions (budgets, focus areas like economic opportunity, mobility and community services as well as technologies like AI and big data analytics). Standardization of applications provided us with a suitable corpus for topic modeling and reduced an "apples-to-oranges" comparison of dissimilar content and media.

3.3.1 Modeling

Data preprocessing is required for modeling. First, we needed to translate three SCC documents from French into English. Canada is officially bilingual and proposals were written in English or French. Regardless of the original language, topic modeling is conducted on Twitter data and tweets are typically translated into English (Bakillah and Liang 2015). Topic modeling appears to work better in English than other languages (Székely and vom Brocke 2017) likely because large language models have been optimized for English. However in Canada, the language of analysis represents a political decision.

Topic modeling requires that we manually predetermine the number of topics, which controls the level of detail of the model (Jacobi et al. 2016). There is no set rule to ascertain a number into which LDA should assign the terms. As introduced in the prior section, determining the number of topics creates a trade-off since "the goal is to describe the data with fewer dimensions (topics) than are actually present, but with enough dimensions so that as little relevant information as possible is lost" (Jacobi et al. 2016, p 93). A greater number of topics lead to better representation of the corpus but with each addition gains less information. This can be measured by a topic coherence score (Röder et al. 2015). We found 20 topics are most suitable to the GAM and 13 topics to the FPM by manually evaluating the topics produced by the model when run for multiple numbers of topics (i.e., every increment from 10 to 30).

3.3.2 Topic Comparison

We were interested in comparing the two corpora to see whether we could infer government preferences in terms of what constitutes smartness (where smartness also could mean political decisions about the geographic distribution of grant finalists and awardees). For this we used the cosine measure, which is frequently applied to evaluate the similarity between corpora (cf., Vulić et al. 2011) and among topics (e.g., He et al. 2009). A cosine measure can be used even if corpora are dissimilar in size because it assesses the angle–the cosine–between one corpus's word vectors and another's to compare how well they align. A similarity score varies from $0-1$. Measures of sufficient similarity (i.e., how close to 1) vary but other authors affirm that 0.5 and above are considered similar (Vulić et al. 2011). We obtained a score of 0.861, which means the two corpora possess highly similar distributions of words.

After gauging the similarity of the corpora, we used the cosine measure to compare the similarity of the topics in the GAM and FPM. The cosine measure analyzed over ten thousand terms to calculate the similarity for each pair of GAM-FPM topics. There is no definitive threshold of similarity scores to determine if two topics are similar or dissimilar. Once again, subjective assessments are necessary to make the call. The similarity of topics was influenced by how distinct the topics were. The greater the number of topics, the greater the specificity of the representation of each topic and the higher the likelihood of a high similarity measure (Aletras and Stevenson 2014; Jacobi et al. 2016). Small corpora (1.5 million words) limit the distinctness (i.e., topic distance) and therefore reduce the prospect of strong "matches". We can still detect partial similarity (as low as 0.25 and as high as 0.70) because, as Jacobi et al. (2016) observes, the measure is detecting word clumping or subclusters within topics. When there is a small value of similarity score, we cannot simply claim that the topics are unrelated. There still can be an opportunity to identify and report comparable subtopics.

3.4 Results

A wide variation in the population sizes of communities submitted SCC applications, from a minimum of 185 residents to a maximum of 2,731,571 residents. The federal government categorized prizes based on population size. One prize of $5 million was open to all communities under 30,000 people; two prizes of $10 million were open to all communities under 50,000 people; and another prize of $50 million was open to communities regardless of population (Infrastructure Canada 2017). Of the 130 original applicants and 20 finalists, ultimately $5 million was awarded to the Town of Bridgewater in the Province of Nova Scotia; $10 million to City of Guelph and County of Wellington in the Province of Ontario and to Nunavut Communities in the Territory of Nunavut and $50 million to City of Montreal in the Province of Quebec.

3.4.1 Modeling of Grant Applications

For the sake of brevity, all our results are presented in Table 3.1, which also compares FPM topics to GAM topics. Column 4 shows 20 different GAM topics–numbers refer to the order in which they appeared in the model. The topic contribution (TC) rate (Column 5) is the word/ term probability, which aggregates the frequency of the over ten thousand terms (we use "term" instead of 'word' or 'token' for consistency) in the corpus divided by the frequency of that topic's terms in the corpus. Word probability normalizes different terms in the corpus and measures the degree to which a topic captures specific terms in the corpus. To some extent, TC indicates the popularity of a topic. In Column 6, we identify the most representative application (MRA) for each topic. The TC to the MRA refers to the sum of the frequency of each term in a topic that belongs to a particular application, divided by the frequency of that term in the application.

We assigned a name to each topic (i.e., cluster of terms) by reviewing the 30 most frequent terms associated with this topic. It was straightforward to assign topic names to most clusters. For example, one cluster of terms with frequent words, including

Table 3.1 Topics of FPM and GAM, including the topic contribution (TC), percentage of the topic to the overall corpus and either the most representative proposal (MRP) or application (MRA) of that topic. The last column shows significant cosine similarities (COS) between the two models. Bolded similarities refer to the highest scores of all GAMs compared to each FPM topic; underlined similarities refer to the highest scores of all FPMs compared to each GAM topic

Topics of FPM	FPM TC %	The MRP	Related topics of GAM	GAM TC %	The MRA	COS
1. Public engagement	15.3	Airdrie area	11. Public engagement	5.3	Stratford	**0.660**
			3. Public consultation	7.4	Mashteuiatsh	0.504
2. Risk and management	15.1	Richmond	8. Management	6.1	Saint-Nazaire	**0.629**
			1. Partnership	10.3	Frog Lake First Nation	0.535
			7. Urban Planning	6.1	Williams Lake	0.435
3. Health care	13.1	Cote Saint-Luc	5. Citizen and urban development	6.9	Coaticook	**0.466**
			19. Health care	1.1	Airdrie area	0.430
			15. Innovation	2.8	Halton Hills	0.308
			13. Demography	3.6	Banff	0.284
			18. Safety and emergency response	2.6	Richmond	0.226
4. Data solutions	9.1	Montreal	10. Data solutions	5.8	Guelph and Wellington	**0.597**
			12. Connected technology	4.9	Cote Saint-Luc	0.574
5. Information privacy	8.3	Kelsey, The Pas, Cree Nation of Opaskwayak	2. Digital services	9.0	Strathcona	**0.518**
6. Youth/child and culture, education	7.7	Biigtigong Nishnaabeg	9. Youth/child	5.9	Nunavut Association of Municipalities	**0.673**
			14. Culture, education and language	2.8	Biigtigong Nishnaabeg	0.253
7. Digital services	6.3	Saskatoon	2. Digital services	9.0	Strathcona	**0.575**
8. Housing and energy	5.5	Bridgewater	17. Housing and energy	2.7	Cree Nation of Eastmain	**0.570**
9. Food and agriculture	5.5	Guelph and Wellington	20. Food and agriculture	0.3	Halifax Regional Municipality	**0.595**
			4. Economic opportunity	7.0	Powell River	0.432
10. Transportation/mobility	4.9	Greater Victoria	6. Transportation/mobility	6.4	Dieppe	**0.568**

(continued)

Table 3.1 (continued)

Topics of FPM	FPM TC %	The MRP	Related topics of GAM	GAM TC %	The MRA	COS
11. Innovation and funding	4.0	Nunavut Association of Municipalities	10. Data solutions	5.8	Guelph and Wellington	**0.339**
12. Housing and culture	3.0	Cree Nation of Eastmain	17. Housing and energy	2.7	Cree Nation of Eastmain	**0.377**
13. Utilities cost	2.2	Tri-council region	16. Utilities cost	2.8	Kelsey, The Pas, Cree Nation of Opaskwayak	**0.381**

"partner", "provide", "partnership", "design", and "development", suggested partnerships that were created to develop smart initiatives. The cluster of terms for connected technology including terms that encapsulated new digital technologies for networking and interoperability in smart cities.

A few instances proved challenging because of seemingly distinct terms. One cluster included "provide", "preliminary", "select", "service", "area", "number", and "letter". To name this topic "innovation", we referred to its MRA, the Town of Halton Hills, ON. The application stated that the town "will become the leading twenty-first century low carbon community by accelerating the adoption of electric vehicles through the development and deployment of a network of internet connected electric car charging stations".

Popular topics, like "partnership", "digital services" and "public engagement" appeared in all 117 applications. Topics like "health care" (59) and "food & agriculture" (17) occurred in the fewest applications. Topic modeling requires careful curation since high performance topics, like "partnership", can be anodyne and therefore add little value to a deeper understanding of topics significant to specific communities (Lim et al. 2018). No community would likely reject the need for partnerships in developing a smart city. By contrast, low performing topics like "culture, education & language" suggested the needs of smaller but important segments of Canadian society.

Whereas topics ranged from "culture, education & language" and "demography" to "utilities", GAM topics revealed a prevalence of "soft", non-explicitly technological themes. There were several generic references to technology ("digital services", "data solutions", "connected tech", "innovation"). Even with ostensibly hard, technology-dependent domains, we found keywords referring to soft domains. For example, food & agriculture contained soft ("farmers", "affirm") and hard elements ("drone", "robotic"). Topic modeling allowed us to distill 117 applications with an abductive approach, finding community requests did not align neatly with the majority of the literature.

3.4.2 Modeling of Finalists' Proposals

Table 3.1's first column shows thirteen topics generated for the FPM with their TC rates listed in the second column and their most representative proposal (MRP) in the third column. Based on TC rates, the most popular topics were "public engagement", "risk & management" and "health care", together capturing 43.5 percent of total terms. The first two were unsurprising because finalists were directed to address both in the proposal. The rest amounted for between two and nine percent. FPM topics were straightforward to label except "housing & culture." Its MRP from the Cree Nation of Eastmain, QC illustrated how housing issues are rooted in the history of Indigenous communities and the legacy of government interventions in that history.

Inclusive of "housing & culture", six of thirteen topics involved multiple subjects. Topic 6 joined culture to education and related them to youth/child. Students in K-12 are foci for education programs that also include cultural subjects in proposals from Indigenous communities. Topic 11 uncovered the importance of funding and budgets to innovation. The increasing proportion of multi-subject topics resulted from FPM's aggregated representation of topics but also implied that the communities viewed issues as interconnected.

Overall, we found a replication of topics from the GAM, with one notable exception of information privacy. We saw topics slightly more on the technical or at least technocratic side, which could be explained by the instructions demanding details on implementation for final proposals.

3.4.3 Comparing the First Round to the Finalists' Round

We compared every pair of GAM-FPM topics (260 comparisons) and found the highest similarity score of each topic in the comparison (the last column of Table 3.1). Ten of thirteen FPM topics' scores are bold and underlined (i.e., most similar to GAM topics and converse). Of note, the similarity of topic names was a result of our human-centered labeling activities, not our statistical analysis. Cosine similarity scores ranged from 0.226−0.673. Cosine similarity showed the degree of alignment of one corpus's topic to another. Unsurprisingly, high levels of alignment were found in public engagement-related topics. A similarity measure also revealed unforeseen alignments and the importance of subclusters. The number three ranked FPM topic "health care" represented a domain that spanned multiple subjects, although statistically its most similar GAM topic was "citizen & urban development". In examining the 30 top terms of related GAM topics, we found occurrences of subclusters related to health care, including citizen/people, quality of life, and technology. Also interesting was health care's ranking near the top of FPM but lower in GAM, perhaps reflecting the choices of the federal government, conscious or subconscious, in selecting finalists. "Information privacy" emerged as a topic in the FPM. When we read the MRA of GAM associated with digital services and examined the top 30 frequent terms, we saw alignments with subclusters related to information and data. Cosine similarities also pointed to instances in which inspection of baseline texts could identify any similarities.

3.5 Discussion

Our topic models help us understand what the smart city has meant to Canadian communities, what the Canadian federal government envisioned was a smart city, and how communities wrote applications that they believed had the best chance of success. We then track the dynamics of the SCC through the similarities among topics. Lastly, we discuss the utility of topic modeling for bottom-up analysis of smart cities.

3.5.1 What Constitutes the Canadian Smart "City"?

Modeling results were consistent with the literature's uncertainty around defining what constitutes a smart city. Instead, Canadian smart city initiatives varied to match specific geographic or local policy needs (Clement and Crutzen 2021), from data and technologies to public participation, and from transportation to health care. The analysis suggests that communities envisioned the smart city through different approaches even as they all asserted their potential to become smart. Where the SCC differed from the literature was the decentering of an urban discourse and the application of the smart city concept to a wide variety of community types.

Numerous topics appeared in both urban and rural communities, which reinforce arguments that rural areas can be as technical and smart as cities (Shearmur et al. 2020). Recalling Neirotti et al. (2014), soft domains like culture, education & language were proposed by rural communities alongside hard domain topics, like housing & energy. Agriculture is highly technologically sophisticated even as it rarely appears in the smart city literature. Food & agriculture was not strictly rural but appeared in both urban and rural applications. The MRA, Halifax Regional Municipality, included both the City of Halifax, the largest city in Nova Scotia and its surrounding rural areas. This may be atypical compared to more uniformly dense communities but suburban communities may encompass rural extents as well. Topics can potentially be used to detect these differences in urban structure.

Compared to other countries, Canada has strong representation by Indigenous communities, many of which are classified as geographically remote. This further broadens the scope of "city" and the topics addressed. The predominant topic from the Biigtigong Nishnaabeg First Nation was culture, education & language, a topic not ostensibly associated with the smart city. In the application, the community reported they developed an education program empowered by technologies like cloud computing, mobile applications, open data, and video analytics. By proposing both

Nishnaabe knowledge and technology, Biigtigong Nishnaabeg communities hoped their youth could attain full language literacy in the endangered Nishnaabe language (grant application from Biigtigong Nishnaabeg). We should ensure that the smart cities do not contribute to a recolonization of Indigenous communities via urban innovations and assumptions.

Topic models revealed regional cultural distinctions. Community applications from the Province of Quebec expressed a distinct cultural identity compared to other Canadian provinces. This difference is primarily manifest with language systems where French, instead of English, is the official language spoken in Quebec (Statistics Canada 2022). The cultural difference was further embedded in "a vision of society that recognizes community participation as a fundamental exercise in citizenship and democracy, and as a means for empowering citizens" (Laforest 2007, p 172). Applications from most Quebec communities clustered keywords into the topic "citizen & urban development" more than the rest of Canada. Place and cultural specificity has been found in other countries (e.g., Catalonia in Spain, Southern US States). Smart city rhetoric can obscure these differences (Sepasgozar et al. 2019).

3.5.2 What Constitutes the Canadian "Smart" City?

The Canadian federal government chose not to define a smart city narrowly but rather characterized it as an approach "to achieve meaningful outcomes for residents by leveraging the fundamental benefits that data and connected technology have to offer" (Infrastructure Canada 2017, p 2). Application guidelines listed multiple visions of a smart city: having residents "feel safe and secure," "move around my community," "enjoy a healthy environment," and "be empowered and included in society" (Infrastructure Canada 2017). The vague and aspirational messages allowed for broad interpretation and contrasted with reminders that proposals should be rooted in connected technology. Interestingly, environmental sustainability (e.g., climate resilience), a current focus of the literature (Tura and Ojanen 2022), did not emerge in SCC proposals.

Applicants were encouraged to mention specific technologies to accomplish their projects. Applications contained numerous references to AI, augmented reality, virtual reality, cloud computing, Internet of Things, mobile applications, and big data analytics. With few exceptions (e.g., agricultural drones), these words failed to emerge as keyword descriptors of topics. Moreover, what constitutes innovative technology in one community did not occur in another. Frog Lake First Nation proposed to map unofficial garbage dump sites with geographic information systems (GIS). GIS is hardly a new technology in municipalities; yet it was wholly new for many small communities in Canada. Frog Lake First Nation appeared to be an innovation "laggard," which suggests communities possessed different understanding of smartness. Because we lack consensus on what constitutes a smart city, a laggard in certain technologies is not necessarily a laggard in the smart city.

Applicants who were selected by the judges as finalists arguably were "smarter" relative to applicants who were not; topics of FPM should therefore represent smarter concepts. Dominant topics (topics with highest contribution rates) for finalists covered a mix of hard domains like data solutions, utilities cost and housing & energy, and soft domains like culture, education & language, and health care. All winning proposals appeared in the list as MRPs of a FPM topic, which means the FPM was able to identify all the SCC winners. Eight of the finalists' applications emerged as MRAs of GAM topics. The GAM identified 40 percent of the finalists and thus indicates what was sufficiently smart to make it to the next stage of the competition. The most representative proposals contained the most prevalent topics across all applications so unfortunately unusual applications were not awarded. Topic modeling does not function as a predictive model but that is what appears to have occurred.

The relative topic contribution rates of FPM topics shown in Table 3.1 also suggested what constitutes smarter. The top three topics, public engagement, risk & management and health care, had a contribution rate of over 13 percent. Since finalists were advised to emphasize public engagement, risk, management, data, technology, and privacy in their final proposal (Infrastructure Canada 2018) we did not count FPM topics 1, 2, 4, and 5 as automatic indicators of smartness. This suggests an emphasis on health care secured a spot as a finalist.

Knowing the topics of the FPM, we could use similarity scores to identify where portions of FPM—the "smarter topics"—appeared in the GAM. Table 3.1 shows GAM topics with the highest similarity to a corresponding FPM topic (bolded scores). Several FPM topics matched two or more GAM topics, although only one of GAM topics held the highest similarity (i.e., bolded scores). The GAM topics with the greatest similarity, that is consisting of sufficiently related subclusters, were relatively smarter. The FPM health care topic illustrates the importance of subclusters. In a country that highly values its national health care system, one might think that applications explicitly linked to health would be privileged; however, health care in the GAM ranked 19 out of 20. Health's most similar GAM topic was citizen & urban development, which indicates a broader categorization to community health that includes quality of life. An example is Nunavut Communities, one of the finalists and winners whose application proposed to improve mental health among the Nunavummiut. We found that cosine similarities can assist in detecting smartness but smartness cannot be looked at in isolation from context.

3.5.3 What Is the Utility of Topic Modelling for Abductive Inquiry About Smart Cities?

Sections 3.5.1 and 3.5.2 reveal uniquely Canadian insights to smart cities; moreover, we find important cultural details from Quebec and from Indigenous communities that do not appear in standard characterizations of smart cities. Should we adopt emerging approaches using topic modeling on long documents (e.g., Nicolas et al. 2021; Clement and Crutzen 2021) we may gain universality and comparability at the expense of context-specific nuance. To illustrate this problem, we replicated the labeling used in Nicolas et al. (2021), based on classifications developed by Giffinger et al. (2007) and Neirotti et al. (2014). Table 3.2 shows the impacts of fitting the topics of GAM into these two pre-existing classification frames. For example, a topic expressing the importance of language preservation is completely absent in Giffinger et al. (2007) or captured as 'knowing foreign language' or 'requirement for multiple languages'. Although considered smart in Giffinger's study, this unfortunately would be seen as insulting to Indigenous communities, since foreign language requirements would likely be their own (e.g., learning Inuktitut in Quebec). Agriculture is missing. Children and elderly are subsumed into other classifications. All advances in technology are missing, not just new advances (robotics) but also relatively old (GIS). The FPM revealed nuances in public participation that are not well-covered in smart governance. Indeed citizen engagement is split across multiple classifications.

Clement and Crutzen (2021) compared corpora through human interpretation of keywords in the clusters but ignored statistical analysis, for example available through a cosine measure. We may advocate for a human-centered approach. However, intervention and abductive analysis must be tempered by the opportunities and constraints of the computational method.

3.6 Conclusion

Three findings arise from our topic modeling and similarity comparison of the SCC. First, topics related to rural, regional and Indigenous communities challenge our notion of the "city" in the smart city. A smart city may be a sparsely populated area. In the future, rural and remote communities may have an increasingly important role in smart city discourse, for example as testbeds that combine hard and soft domains. Regardless of rural–urban context, a country will invariably contain multiple cultural and geographic identities, which challenge the universality of smart city initiatives.

Second, a variety of measures, many of which are context-dependent, can be used to infer what constitutes smarter. Certain domains may be irrelevant to the needs of a specific community; they may possess neither the capacity nor the skills to forcefit a particular domain to a community need. We do not want to label communities as less smart than communities with greater resources. Those measures include the predominance of soft themes, further demonstrating that smartness does not equal 'tech'. Increasingly, the distinction between soft and hard domains represents a false choice in the smart city. Data and connected technology permeate every aspect of cities and it is problematic to create an ordinality that prioritizes hard domains over soft domains. The Canadian smart city appears to have moved beyond the technology-driven phase toward technology-enabled, city-led and citizen co-creation phases.

Lastly, our approach shows the value of a bottom-up abductive inquiry about the smart city over conventional methods, including identifying 40% of the finalists

Table 3.2 Categorization of GAM topics according to Giffinger et al.'s and Neirotti et al.'s classifications

Topic names	Giffinger et al. (2007) categories	Neirotti et al. (2014) categories
Partnership	Smart governance (participation) Smart economy (competitiveness)	Soft
Digital services	Smart mobility (transportation and ICT)	Hard
Public consultation	Smart governance (participation) Smart people (social and human capital)	Soft
Economic opportunity	Smart economy (competitiveness)	Soft
Citizen and urban development	Smart people (social and human capital) Smart living (quality of life)	Soft
Transportation/mobility	Smart mobility (transportation and ICT)	Hard
Urban Planning	Smart governance (participation)	Soft
Management	Smart governance (participation)	Soft
Youth/child	Smart living (quality of life) Smart people (social and human capital)	Soft
Data solutions	Smart mobility (transportation and ICT) Smart living (quality of life)	Hard
Public engagement	Smart governance (participation) Smart people (social and human capital)	Soft
Connected technology	Smart mobility (transportation and ICT) Smart living (quality of life)	Hard
Demography	Smart people (social and human capital)	Soft
Culture, education and language	Smart living (quality of life)	Hard

(continued)

Table 3.2 (continued)

Topic names	Giffinger et al. (2007) categories	Neirotti et al. (2014) categories
Innovation	Smart economy (competitiveness) Smart mobility (transportation and ICT)	Hard
Utilities cost	Smart environment (natural resources)	Hard
Housing and energy	Smart living (quality of life) Smart environment (natural resources)	Hard
Safety and emergency response	Smart living (quality of life)	Hard
Health care	Smart living (quality of life)	Hard
Food and agriculture	–	Hard

and the four winning cities. Our approach offers a useful albeit imperfect inference of what the federal government considered smart. Did the federal government pre-determine the smart topics (e.g., health)? Were there political considerations related to the geographic distribution of awardees? Topic modeling is not ideal to quantify the political calculus of government but it does provide insightful analysis of large amounts of qualitative text.

Topic modeling is useful for a wide variety of urban research, for example master plans and large-scale public feedback sessions (e.g., from transportation planning). Our research taught us that we should never rely on technology alone, whether that is algorithm, data or hardware, so topic modeling should be embedded in a mixed method combined with qualitative analysis. That sounds obvious but algorithmic evangelism coupled with smart city hardware (e.g., Internet of Things sensors) can push us to technology solutionism approaches (Morozov 2013), especially as algorithms are optimized and data is bigger. Our methods can accompany Cohen's (2015) observation about the evolution of smart cities from high-tech and towards human-centered approaches.

Acknowledgements The authors wish to acknowledge the funding for this research was provided by the Social Sciences and Humanities Research Council of Canada SSHRC 895-2012-1023.

References

Albino V, Berardi U, Dangelico RM (2015) Smart cities: Definitions, dimensions, performance, and initiatives. J Urban Technol 22(1):3–21

Aletras N, Stevenson M (2014) Measuring the similarity between automatically generated topics. In: Wintner S, Goldwater S, Riezler S (ed) Proceedings of the 14th conference of the European chapter of the association for computational linguistics. ACL, Gothenburg, Sweden, pp 22–27

Appio FP, Lima M, Paroutis S (2019) Understanding smart cities: innovation ecosystems, technological advancements, and societal challenges. Technol Forecast Soc 142:1–14
Bakillah M, Li RY, Liang SHL (2015) Geo-located community detection in twitter with enhanced fast-greedy optimization of modularity: the case study of typhoon Haiyan. Int J Geogr Inf Sci 29(2):258–279
Batty M (2013) Big data, smart cities and city planning. Dialogues Hum Geogr 3(3):274–279
Caragliu A, Del Bo C, Nijkamp P (2011) Smart cities in Europe. J Urban Technol 18(2):65–82
Ching TY, Ferreira J (2015) Smart cities: Concepts, perceptions and lessons for planners. In: Geertman S, Ferreira J, Goodspeed R, Stillwell R (eds) Planning support systems and smart cities. Springer, Cham, pp 145–168
Clement J, Crutzen N (2021) How local policy priorities set the smart city agenda. Technol Forecast Soc 171:120985
Cocchia A (2014) Smart and digital city: A systematic literature review. In: Dameri RP, Rosenthal-Sabroux C (eds) Smart city. Springer, Cham, pp 13–43
Cohen B (2015) The 3 generations of smart cities: inside the development of the technology driven city. Fast Company. https://www.fastcompany.com/3047795/the-3-generations-of-smart-cities. Accessed 4 Feb 2023.
Crowston K, Allen EE, Heckman R (2012) Using natural language processing technology for qualitative data analysis. Int J Soc Res Method 15(6):523–543
Dameri RP (2013) Searching for smart city definition: a comprehensive proposal. Int J Computers Technol 11(5):2544–2551
Debnath R, Bardhan R (2020) India nudges to contain COVID-19 pandemic: A reactive public policy analysis using machine-learning based topic modelling. *PLoS One* 15(9). https://doi.org/10.1371/journ al.pone.0238972
Gan J, Qi Y (2021) Selection of the optimal number of topics for LDA topic model—taking patent policy analysis as an example. Entropy (basel) 23(10):1301
Giffinger R, Fertner C, Kramar H, Kalasek R, Pichler-Milanović N, Meijers E (2007) Smart cities–ranking of European medium-sized cities. Final Report. Centre of Regional Science at the Vienna University of Technology. https://www.smart-cities.eu/download/smart_cities_final_report.pdf. Access 30 Jan 2023
Han H, Hawken S (2018) Introduction: Innovation and identity in next-generation smart cities. City, Culture Society 12:1–4
He Q, Chen B, Pei J, Qiu B, Mitra P, Giles L (2009) Detecting topic evolution in scientific literature: how can citations help? In: Proceedings of the 18th ACM Conference on Information and Knowledge Management, Hong Kong, China, pp 957–966
Infrastructure Canada (2017) Smart cities challenge applicant guide. https://impact.canada.ca/sites/default/files/2017-11/SCC_Applicant_Guide.pdf. Accessed 18 May 2019.
Infrastructure Canada (2018) Smart cities challenge finalist guide. https://impact.canada.ca/en/challenges/smart-cities/finalist-guide. Accessed 18 May 2019.
Jacobi C, Van Atteveldt W, Welbers K (2016) Quantitative analysis of large amounts of journalistic texts using topic modelling. Digit J 4(1):89–106
Javed AR, Shahzad F, ur Rehman S, Zikria YB, Razzak I, Jalil Z, Xu G, (2022) Future smart cities: requirements, emerging technologies, applications, challenges, and future aspects. Cities 129:103794
Johnson PA, Acedo A, Robinson PJ (2020) Canadian smart cities: Are we wiring new citizen-local government interactions? Can Geogr 64(3):402–415
Laforest R (2007) State and community sector relations: crisis and challenges in Quebec. Philanthr 20(3):171–184
Lim C, Kim KJ, Maglio PP (2018) Smart cities with big data: Reference models, challenges, and considerations. Cities 82:86–99
Morozov E (2013) To save everything, click here: the folly of technological solutionism. Public Affairs, New York

Mosannenzadeh F, Vettorato D (2014) Defining smart city. A conceptual framework based on keyword analysis. J Land Use, Mobil Environ. https://doi.org/10.6092/1970-9870/2523
Neirotti P, De Marco A, Cagliano AC, Mangano G, Scorrano F (2014) Current trends in smart city initiatives: Some stylised facts. Cities 38:25–36
Nicolas C, Kim J, Chi S (2021) Natural language processing-based characterization of top-down communication in smart cities for enhancing citizen alignment. Sustain Cities Soc 66:102674
Papadopoulos T, Charalabidis Y (2020) What do governments plan in the field of artificial intelligence? analysing national AI strategies using NLP. In: Charalabidis Y, Cunha MA, Sarantis D (ed) Proceedings of the 13th international conference on theory and practice of electronic governance. ACM, New York, pp 100–111.
Pollitt C, Hupe P (2011) Talking about government: the role of magic concepts. Public Manag Rev 13(5):641–658
Röder M, Both A, Hinneburg A (2015) Exploring the space of topic coherence measures. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. ACM, New York, pp 399–408.
Rosen J, Alvarez León LF (2022) The digital growth machine: urban change and the ideology of technology. Ann Am Assoc Geogr 112(8):2248–2265
Sepasgozar SM, Hawken S, Sargolzaei S, Foroozanfa M (2019) Implementing citizen centric technology in developing smart cities: a model for predicting the acceptance of urban technologies. Technol Forecast Soc 142:105–116
Shearmur R, Charron M, Pajevic F (2020) Pourquoi seules les villes sont elles qualifiées d'intelligentes? Un vocabulaire du biais urbain. Can Geogr 64(2):310–322
Shelton T, Zook M, Wiig A (2015) The 'actually existing smart city.' Camb J Reg Econ Soc 8(1):13–25
Spicer Z, Goodman N, Olmstead N (2021) The frontier of digital opportunity: smart city implementation in small, rural and remote communities in Canada. Urban Stud 58(3):535–558
Statistics Canada (2022) Mother tongue by geography, 2021 Census https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/dv-vd/language-langue/index-en.html. Accessed 5 Feb 2023
Székely N, vom Brocke J (2017) What can we learn from corporate sustainability reporting? deriving propositions for research and practice from over 9,500 corporate sustainability reports published between 1999 and 2015 using topic modelling technique. PLoS ONE. https://doi.org/10.1371/journal.pone.0174807
Tura N, Ojanen V (2022) Sustainability-oriented innovations in smart cities: a systematic review and emerging themes. Cities 126:103716
Vulić I, De Smet W, Moens MF (2011) Identifying word translations from comparable corpora using latent topic models. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. ACL, Stroudsburg, PA, pp 479–484.
Zheng Z, Sieber RE (2022) Putting humans back in the loop of machine learning in Canadian smart cities. T GIS 26(1):8–24

Chapter 4
The Venue Code: Digital Surveillance, Spatial (Re)organization, and Infrastructural Power During the Covid Pandemic in China

Xiaoling Chen

Abstract China has insisted on a zero-Covid policy for almost three years, deploying many disciplinary containment measures and surveillance tools. This chapter examines a surveillance tool called the "venue code" which is claimed to improve contact tracing of Covid outbreaks. My research is based on ethnography and discourse analysis of online textual and visual materials, capturing the integration of the venue code into China's Covid containment regime. I analyze four aspects of the practice: bureaucratic structures, integration, spatial (re)organization, and technologies. I argue that the venue code mediates residents' experience of cities and restricts urban mobility through the introduction of physical and digital infrastructure, while granting local actors more power. By doing so, this study draws attention to the ways in which the responsibility for achieving the zero Covid objective was shifted onto local actors and Chinese citizens, and state presence in communities and despotic power were strengthened.

Keywords Venue code · Digital surveillance · Infrastructure · Spatial organization · Pandemic

> Teacher Lu Renbing, an instructor in University X in Beijing, was living in the university residential community adjacent to the campus. Two days after her last Covid test, she checked the operation hours of the four testing kiosks within walking distances on a smartphone app—two inside the gated community, two on public spaces outside—and picked the one along the Moon River with 5-minute electric scooter ride. She got in the line of several residents, tapped her Identification Card on the ID reader that automatically read her national ID information in the first window, and followed the line to the next window to get her saliva sampled. The whole process took her roughly 15–20 minutes each time. Upon her return to the community, one of three guards at the gate indicated that she should stop and scan a QR code (the changsuo ma). She took out her phone and turned on a scanner within the Beijing health app. She scanned the code, and her phone read out loud: "two days since Covid test, green code, pass" after a green code appeared. She showed this green code to the guard. The guard nodded his head and permitted her pass into her community.

X. Chen (✉)
Department of Geography, University of Colorado Boulder, Boulder, CO, USA
e-mail: xiaoling.chen@colorado.edu

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_4

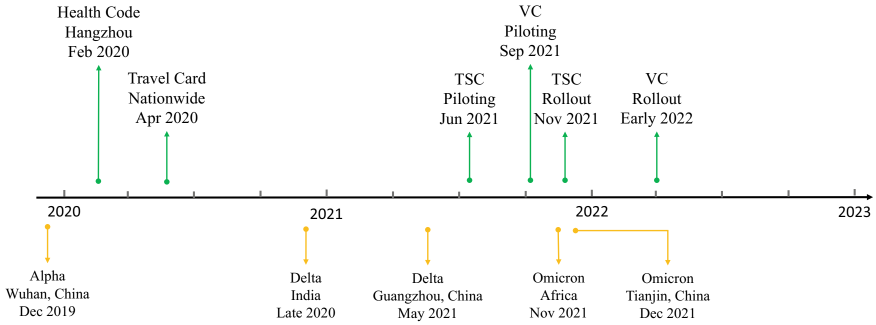


Fig. 4.1 Emergence of Covid variants and integration of digital surveillance in China. *Note* TSC stands for the Time–Space Companion project (Chen and Oakes 2023) and VC stands for the venue code (this chapter)

The routine of Teacher Lu illustrated some ways in which people's city experiences within cities was mediated by big-data technologies such as a health app and the venue code—along with other related containment measures like regular Covid tests. On December 7, 2022, China lifted its zero-Covid policy, resulting in the rapid collapse of a complex Covid containment apparatus established over the past three years. This apparatus featured mass testing, lockdowns, quarantine camps, and digital surveillance, the strengthening and expansion of which were associated with the emergence and mutation of the Covid-19 virus. Figure 4.1 illustrates a positive association between the emergence of different Covid variants and the introduction of four types of digital projects: the health code (Liang 2020; Zhang 2022; Zhou et al. 2021), the travel card, the time–space companion project (Chen and Oakes 2023), and the venue code (this article). By introducing these projects, Chinese citizens' time, space, and bodies were increasingly subject to social management in exchange for the health and safety promised by the government. While the Chinese government granted more and more power to community-based grassroot actors to achieve its zero Covid objectives, the fact that these objectives became increasingly unattainable motivated these actors to resort to excessive, draconian measures. Media headlines have shown how containment measures had caused many social, economic, and political costs, raising the questions of why Chinese citizens still complied and how these costs were produced. This chapter attempts to answer these questions by closely examining a contact tracing project called the venue code. The work outlines the ways in which people's movements were monitored, controlled, and restricted and thus mobility and behaviors were changed.

The term *changsuo* ma (场所码), or venue code, was shown as a QR code generated as an identifier of a specific location to the finest level allowed such as the North Gate of University X, and ABC grocery store—University X branch. When a venue code was scanned, residents' information including the location, the time they enter the location, and the health code information (i.e. Covid test status, vaccination status, and health code color) would be recorded, along with unique personal information already documented in their health app: name, national identification number, phone number, date of birth, and ID card issuance institute, etc. It was claimed that by integrating this practice into the health code system, the health authorities could verify residents' Covid status, instantaneously document their footprint to realize timely, targeted contact tracing and curb Covid spreads, while effectively monitoring mobility and gathering patterns as part of the normalization of Covid prevention while protecting privacy (Datong Health Commission 2022). By June 2022, the venue code was already widely implemented as a significant contact tracing strategy in almost all urban areas. For example, Suzhou City began to use the venue code from March 27, 2022, and by April 7, its Big Data Bureau (*da shuju ju*) of the city recorded use of 1.7 million person-scans (Suzhou Daily 2022). Given the amount of data such a tool can generate, and the potential impacts of its enforcement on people's life, the limited media reports and intellectual discussion of the project has been surprising.

Much of the discussion on digital surveillance projects for Covid examines the information management and its legal and moral implications (Liang 2020; Stevens and Haines 2020; Zhang 2022; Zhou et al. 2021). This paper does not aim to understand the technical aspects of the venue code per se, for example, how the collected data was managed and processed to trace Covid contacts. Instead, it explores the implications of implementing the tool on people's behavior in urban spaces. Soon after the venue code was widely deployed, many urban residents were indiscriminately subject to intense and frequent lockdowns, camp quarantine, and/or regular Covid testing, casting doubts on the imperative and efficacy of the venue code. I argue that the venue code exemplifies ways in which Covid outbreaks provided opportunities for the Chinese party-state to introduce and expand digital enclosures as part of the Covid containment regime, penetrating the public and private spaces of everyday life.

China's bureaucratic system enabled the party-state to mobilize local actors and Chinese citizens through job insecurity, incentives, and punishments to utilize and transform public and private infrastructures. Fearing administrative punishment and believing in the claimed protection of digital tools, many urban residents were incentivized to change their behaviors to meet the disciplinary objectives embedded within these infrastructures. These behaviors included, but were not limited to, changing urban mobility patterns and limiting access to urban spaces. That is, the venue code epitomizes a process in which the responsibility for containing Covid spreads was gradually shifted from the central and local governments onto grassroot actors and individual citizens. In this sense, the visibility of physical and technological infrastructure on the rural and urban landscapes gave residents a paradoxical sense of both fear and security. That is, before the zero-Covid policy was lifted, most citizens perceived both pandemic risks and their level of safety primarily through the containment measures. The venue code exemplifies how the central and local states can infiltrate civic society and achieve national surveillance goals by deploying infrastructural power for despotic purposes in times of crisis. This further legitimizes state efforts to heighten spatial restrictions and segregation despite observed dissent.

Using case studies, I discuss and compare how this contact tracing tool was implemented in Ling County—a semi-urban space with large rural hinterland and rural pockets in urban areas—and Beijing—a mega urban city. Highly urban areas with prevalent gated communities and centralized business centers where entrances and venues can be clearly defined, facilitate rapid enforcement of surveillance. Whenever needed, I supplement the two case studies with evidence from other cities. In what follows, I first provide a theoretical framework on digital surveillance (Sect. 4.1) and research methodology (Sect. 4.2). Afterwards, I provide detailed evidence and discussions to support my arguments (Sects. 4.3–4.6) and conclude with several profound implications of big-data technologies.

4.1 Digital Surveillance and Infrustructures in China

Surveillance is a core function of modern states in both democratic and authoritative societies (Giddens 1987) and for public health from an epidemiological perspective (Thacker and Berkelman 1988). During the Covid pandemic, it was common for governments to deploy mobile apps as surveillance tools for purposes such as case and death registries, and contact tracing, and to facilitate public health intervention (Zhou et al. 2021). China was the among the countries that implemented the most pervasive, invasive digital surveillance in the name of Covid containment through its well-known Health Code system (Liang 2020; Zhang 2022; Zhou et al. 2021) and three less discussed projects: the travel card, Time–Space Companion contact tracing project (Chen and Oakes 2023), and venue code (this chapter). Most studies on China's digital surveillance focus on a nation-wide surveillance camera system with facial recognition technologies (Batke and Ohlberg 2020; Su, Xu and Cao 2021). This monitoring system builds on, supplements, and expands digitalization projects such as Skynet, Project Sharp Eyes, and Smart City/Campus (Erie and Streinz 2021). While these surveillance cameras were originally deployed to curb and solve crime, there is a heightened state interest to expand them into a broader, murkier realm of social control (Chin and Lin 2022). This could be exemplified by their incorporation into China's social credit system (Bach 2019; Yan 2019; Lam 2021) and grid governance (Chen and Greitens 2022; Gao and Cartier 2022; Mittelstaedt 2021; Tang 2020). The expansion and integration of various projects and systems reflect state concerns on the development of civic engagement and social media platforms which pose significant challenges for social control (Batke and Ohlberg 2020; Xu 2017; Yang and Wang 2021).

The incorporation of big-data technologies that feature algorithmic governance, into the surveillance regime, enhanced panoptic monitoring (Gandy 1993). The most well-known online content censorship program, "the Great Firewall", has blocked Chinese people from accessing critical information and thwarted efforts for civic engagement (Hoffman 2019; Wang 2020a, b). Some studies, however, dismiss the mystery of big data technologies and algorithms, and highlight the intensive manual labor needed for data harvesting, management, and treatment. They are also critical about the social inequalities in terms of socioeconomic status and gender, which the digital economy has taken advantage of, produced, and reinforced (Chen and Oakes 2023; Lukács 2020). Increased interest lies in the *infrastructurization* of state power as private digital infrastructures are deployed to collect, manage, and process personal information, to enhance despotic power (Mann 1984; Tarrow 2018). In this sense, it is not surprising that digital platforms are at times regarded as the fifth utility after water, electricity, gas, and telephone (Chen and Qiu 2019). The state imperative to develop a mass surveillance infrastructure is associated with the emergence of a digital economy featuring large telecom companies (i.e. Huawei, ZTE, Tencent, and Alibaba) and digital platforms (WeChat, Weibo, Didi, and Alipay) in China (Qiu and Bu 2013; Huang and Tsai 2022). Many commentators and politicians are worried that they will take on active roles in strengthening both the infrastructural and despotic power of the Chinese state, within and beyond its territory (Bartz and Alper 2022; Chen et al. 2018).

While the public–private partnership is a common practice in many countries, in the United States, tech giants such as Apple and Google are well protected by laws and state power is thus relatively restrained (Zuboff 2019). On the contrary, in China, the party-state has significant control over the private tech giants, and government regulation of critical technologies such as algorithms are not uncommon (Evans 2023; Huang and Tsai 2022; Kharpal 2022). Censorship and the big-data driven surveillance facilitated by the information and communication sector thus allow the party-state to centralize information and arguably become the sole arbiter of "truth", social trust, and Covid containment. The party-state has succeeded in containing Covid-19 and gained legitimacy and public trust in 2020 and 2021 (Wen 2021). However, its legitimacy was constantly challenged by the highly transmissible and elusive Delta and Omicron variants at the end of 2021, which provided the state with opportunities to roll out more restrictive measures and invasive contact tracing tools (i.e. time–space companion project and the venue code). These were achieved through active experimentation of a containment regime that conceived physical social spaces and everyday activities as potential sites of transmission, and human bodies as possible viral carriers, thus criminalizing those entities and individuals who became aspects of the chain of infection (Chen and Oakes 2023). Yet the fact that the party-state must mobilize all possible resources to curb Covid outbreaks, and the unprecedented restrictions that Covid containment has imposed on the population precisely reflected the party-state's struggle to maintain its legitimacy, and the limitations of its power derived from successful containment.

In particular, the Covid-19 pandemic amplifies social ruptures on a contested front of what a "healthy" society should look like: what was the *truth* about Covid-19, what were the appropriate containment methods, and what parties should be responsible for what and to what extent. Despite pervasive censorship and oppression, Chinese citizens have created and transformed digital spaces, documented facts, collectivized memories, and orchestrated civic engagement on social media (Chen 2020; Zhong and Zhou 2021; Yang 2021). As Warf (2011) has put it: "Chinese censorship and its resistance thus form a continually change[d] front of strategies and tactics". Such a state-citizen dialectic finds its parallels in physical spaces and the everyday life of Chinese people. In her study of urban development through Fuzhou's Lanes and Alleys projects, Chu (2014) illustrates "how infrastructures operate not only in support of state projects of legibility but also to condition some surprising political sensibilities.” In this chapter, I find that while the physical and technological infrastructures associated with the venue code have reinforced state despotic power, they also mediated and revealed how residents understood Covid outbreaks and approached grassroot containment. Chinese people’s attempts to break and breach these infrastructures reflect not only social fear but also public dissent against restrictive state control.

4.2 Methodology

This chapter is based on ethnographic research in the author’s dissertation fieldsite, Ling County, Shanhe City (pseudonyms), Guangdong Province between July 2021 and June 2022, and Beijing during June and July 2022, and three-years of (social) media observational study. Ling County has over 1 million residents with roughly half living the county seat with two street offices, another half living in 12 townships. I served as a visiting scholar in the Ling County People’s Hospital—the largest public hospital in the county—and as a volunteer in many public health facilities (i.e. vaccination sites, local public health institutions, and mass testing sites). I was involved in a wide range of Covid-related activities including but not limited to saliva sampling, maintaining order for Covid-related events, and making phone calls to potential Covid contacts. As a partial “insider”, I performed participant observation in spaces of health governance as a researcher (Livingston 2012; Van der Geest and Finkler 2004), and in rural and urban spaces as a resident (Herbert 2000). I conducted semi-structured interviews with hospital and public health administrators, doctors and nurses, local cadres, volunteers and security guards (Berg 2009).

Interviews were audio recorded upon receiving permission from the participants and transcribed afterwards; audio recording were deleted immediately to protect the participants’ identities after the transcription was completed. Fieldnotes were produced to document detailed observation and field analysis. I performed discourse analysis based on (social) media data on Covid response (Lees 2004; Neuendorf 2017). All photos showed in this chapter were taken by the author and their identifiable information were blurred to maintain the anonymity of the places and interviewees. Whenever needed, I used pseudonyms for places and interviewees to protect their privacy and confidentiality. I also deliberately conceal most online sources for my data. Most of the critical messages were deleted either by the authors themselves or by the social media platforms in which these messages were published. Hence the links I documented become meaningless. I would not reveal netizens’ online pseudonyms or official accounts either to avoid drawing undesired attention to them. These individuals and media bore the stakes of sending messages that were increasingly considered politically sensitive despite their sheer factual nature.

I was, and am being trained, in disciplines of Medical Anthropology, Public Health Sciences, and Human Geography in the United States, and had experiences of the Covid-19 outbreaks and containment policies in China and the United States. As a Chinese woman and a first-generation college student from a low-income family, I have embodied and understand numerous barriers for regular Chinese people to achieve a fulfilling life. During my fieldwork, I was deeply touched and inspired by many Chinese people who strived to make the country a better place. I feel imperative to draw academic attention to ways in which Chinese people's everyday life was mediated by the Covid-19 containment regime in particular, and state policies in general. I hope that this chapter will contribute to the policy-making process in China, and help Chinese people to think through their everyday experiences, struggles, and aspirations.

4.3 Organizational Structure

Enforcement of the venue code was made possible by the power China's intricate hierarchical bureaucratic structure—often called Tiao-Kuai system—possesses and its ability to mobilize its grassroot entities. With some territorial variabilities, Figs. 4.2 and 4.3 exemplifies two common bureaucratic structures using the cases of Beijing (Fig. 4.2) and Ling County (Fig. 4.3), graphing how the Tiao-Kuai structure (on the right) corresponds to China's territorial administrative division (on the left). Tiao means vertical strips, while Kuai means horizontal chunks. Governments provide funding to the functional institutions: the Tiaos of the Health Commission and the CDC correspond to the functions of healthcare and public health, along with the overlapping and contradictory goals intrinsic to such an organizational structure (Lieberthal and Oksenberg 2020). While it is assumed that lower levels of functional institutes also receive funding from their supervising institutes, this amount of money is often much lower than that granted from their horizontal supervisors—governments within the same chunk. Thus, the governments usually determine how the functional institutes should act. To curb Covid spreads, Covid leadership groups, consisting of representatives from corresponding levels of government and their various functional affiliations (i.e. healthcare and public health facilities, law enforcement, industry and commerce), were founded and directly led by the government, forming an extra strip and hierarchal chunks.

The Tiao-Kuai system, upholding the State's zero-Covid policy and passing down orders hierarchically through the leadership groups as "political tasks" (*zhengzhi renwu*) to the lowest rungs, mobilize local officials, public servants and grassroot participants using incentives such as promotion and wage increase, community party membership, and *bianzhi*; the last changes an employment status from a contracted worker to a privileged lifelong public servant. These processes of mobilization, responsiblization, and incentivization resulted in rapid expansion and politicization of grassroot organizations over the last three years. Lying in the heart of the grassroot Covid containment regime were the village and urban neighborhood/residential committees (Tomba 2014; Zhang 2010), grid governance system (*wangge*), and Homeowners' Associations (*yezhu weiyuanhui*) (Wang et al. 2012; Tang 2018). These community-level grassroot organizations had no administrative power; rather, they were elected by the residents and supposed to represent the interests and autonomous

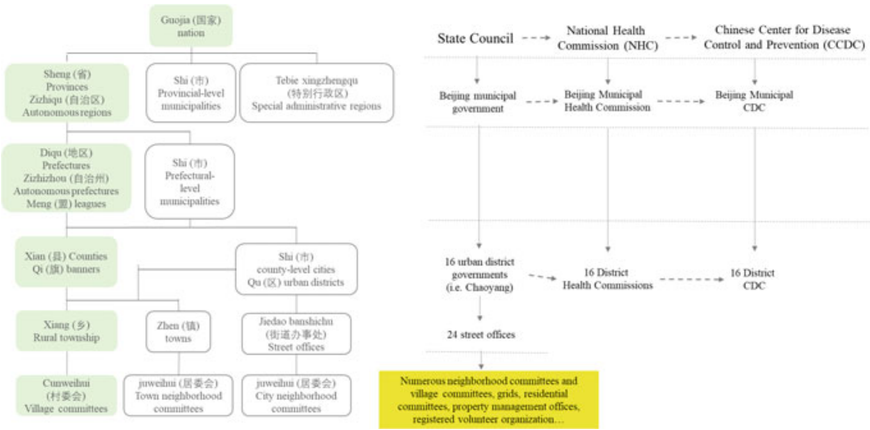


Fig. 4.2 The Tiao-Kuai system of Beijing

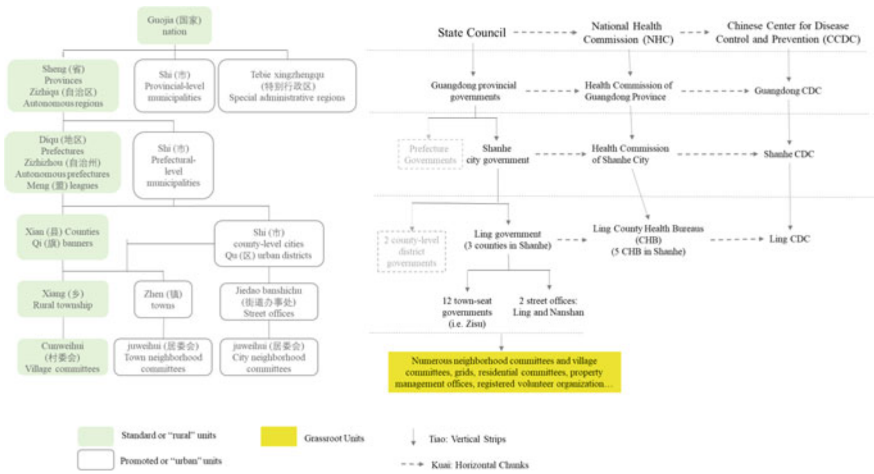


Fig. 4.3 The Tiao-Kuai System of Ling County, Shanhe City, Guangdong Province. *Note* These two figures were produced by the author with a reference to Lieberthal and Oksenberg (2020). Directions of the arrows indicate the hierarchy of the organizational structure. The institutions on the left and at the top are leading those on the right and below them

power of Chinese citizens. However, during the pandemic, they were highly motivated and mobilized by potential economic rewards, sociopolitical status, and political priorities and ideologies, and served as the local enforcers of restrictive state Covid measurements, the effects and results of which at times contradicted their mission. In Ling and Beijing, these grassroot organizations felt an urgent need to enforce the containment measures, and monitored the public and private entities and residents within their purview. Many feared that they would lose their jobs if they were not following the government orders. Such a fear was exacerbated by the unprecedented high unemployment rates in China due to an economic downturn worsened by the zero-Covid policy.

In addition, another type of grassroot organization, officially registered volunteers, were recruited to make up for manpower shortages. Receiving little to no compensation, volunteers were assigned to posts in vaccination sites, mass testing sites, and local public health offices. Their primary responsibilities included but were not limited to maintaining public order, educational campaigns, public mobilization, and contact tracing. Two volunteering jobs I engaged were directly related to the venue code. The first occurred in the Zisu Town Covid-19 Prevention and Control Office with transportation compensation and two daily meals in March 2022. One of my responsibilities was to make phone calls to lists of business operations and asked whether they had deployed a venue code, and take notes on their status (i.e. yes, no, or business close down). My supervisor would then pass this list to the local Bureau of Industry and Commerce for further processing. According to my interviews and participant observation, businesses were persuaded and coerced by local authorities to enforce the venue code at their venues; otherwise, they would face administrative penalties, monetary fines, and/or suspension. The second task I engaged in was to raise awareness about the significance of masks and the venue codes in Covid control. The high unemployment rate provided many ready volunteers. By March 2022, the organization I registered in had a total of 8,563 members, a steep increase of 81% compared to March 2021 (4,726 members).

The increased power of the grassroot organizations was a direct result of the endorsement of the Central government and Chinese Community Party. Given the significant roles local agents have taken during the first year of Covid containment, the Party and the State Council jointly vowed "to strengthen the authority of rural townships and urban subdistricts and place them under more direct leadership of the party" (Ye 2021; CPCC and the State Council 2021). The grassroot units at the community level took on both administrative and service roles for the Chinese people and thus increased the latter's dependance on the former in everyday life. As a local official at a Street Office in Guangdong has confessed, her street office has only five to six hundred people being responsible for serving some hundred thousand residents living in nine communities, "the containment has crushed (*yakua*) many grassroot community workers." (Wang 2023). Rounds of mass testing and strict digital surveillance provided residents a sense of security within their own communities. This gave residents a sense of certainty and security frequent travelers lacked (i.e. truck drivers, see Chen and Oakes 2023) which in turn heightened their hesitance to travel out of their safe zones. Local governments and public health officials also increased compliance by criminalizing daily life and violation of Covid containment measures, further disciplining their residents (Chen and Oakes 2023).

However, having more power also emboldened the grassroot enforcers to exert power in ways that reminded many of the Red Guards during the Cultural Revolution (Yuan 2022). Indeed, draconian measures such as prolonged confinement, forced quarantine, and physical and verbal violence were rampant and became hotbeds of local-citizen disputes, suffering, and deaths. These measures were a result of constant negotiation between the local actors and the residents in the face of an unattainable zero-Covid goal since late 2021. Much effort was put on the part of the local officials and grassroot entities to improve their own capacity to manage the population and to discipline the residents to behave in ways they desired (Sects. 4.4–4.6). Grassroots' deepened involvement in the containment regime was concerning. Many did not think that they had become complicit in causing suffering, tragedies, and a loss of lives which contradicted their stated belief that "People and Their Lives Come First" (*renmin zhishang, shengming zhishang*). Local officials often ascribed suffering and deaths which resulted from excessive containment measures to the managerial incompetence of individual leaders, and stated that "local officials have differential management capacities, and it takes time for them to improve such capacities." Such an understanding implied necessary human sacrifice for a larger goal of public health. It also highlighted a systematic indifference within the bureaucratic structure (Herzfeld 1993) and the tendency of bureaucrats to evade responsibility by strictly implementing task-work (Lieberthal and Oksenberg 2020) which at times contradicted the shifting needs of Chinese people, who required flexibility in the face of an evolving pandemic.

4.4 Definition, Piloting, and Incorporation

The venue code exemplifies ways in which the government attempted to shift public health responsibility from the state to society as it transformed spaces of everyday life into sites for health governance, where intense scrutiny could be imposed upon social actors and individuals. The venue code was designed to replace the manual logbook system for contact tracing, which had been loosely implemented since early 2020 primarily in public institutes such as banks and campuses. By digitalizing the logbook, the government expanded its use in all "public venues" (*gonggong changsuo*) as *critical* sites for Covid containment. According to official classification, these critical sites include shopping malls and every store they hosted, public transit (i.e. entrances and buses), business, and residential communities and their individual buildings. That is, all venues including publicly and privately owned spaces in legal terms, except for private homes were considered *public* venues, and venue owners were compelled to enforce the venue code to avoid potential administrative punishments. Originally, China's contact tracing relied primarily on a self-reporting system aided with scattered urban and business infrastructures such as CCTV and online payment histories, requiring public health officials and law enforcement to conduct lengthy investigation and verify information. With the venue code, real-time information of Chinese citizens' movements with unique identifiers was harvested and centralized, ready for data mining, analysis, and potential manipulation.

Piloting was used to make the practice logistically streamlined, technologically viable, and comparable with other Covid surveillance projects, and to test public acceptance. A few months before the venue code officially appeared in media, it was experimented as an individual applet in Ling County's vaccination sites beginning on September 14, 2021. This process was enabled by a large amount of manual labor from the local actors and residents. As a volunteer, I was asked to introduce to the residents that the code "uses big data technology to guarantee your safety should there be an outbreak". Residents needed to display three separate codes to obtain access to the vaccination sites and to receive a vaccine: the health code, a travel card (Fig. 4.4), and a venue code. At the time, residents complained about the difficulty of navigating three separate applets and the long queues that resulted; it could take each resident up to two hours before they got their vaccine shots. As a result, security guards and volunteers fell prey to complaints while they were also busy solving the technical issues of these applets installed on different smartphone models while assisting in handling vaccination-related paperwork. By the end of 2021, both the travel card and venue code were integrated into the health code applet and could be easily displayed with at most one scan and one click (Fig. 4.5). This integration significantly streamlined the code displaying process when residents requested access to public venues. However, such an integration also risked masking the amount and nature of information that people relinquished. Many experienced policy-exhaustion and became numb and tired of negotiating every step of their daily lives. Some residents confessed that they did not know the purposes of the QR codes and just wanted to get it over with while scanning them. Others ignored the policies or figured out ways to counterfeit or evade them.

Technically speaking, the full incorporation of the venue code into the containment regime proved to be much more difficult than its physical integration into the

Fig. 4.4 A travel card applet (left) and a travel card after entering a phone number (right) (Screenshots from the interviewees)

Fig. 4.5 A regular health code (left) and a health code with granted venue access (right) (Screenshots from the interviewees)

health app. First, it required the full commitment of all public and private entities to deploy this tool at their venues, and interdepartmental coordination and collaboration between the local Administration of Industry and Commerce, and local governments. At least in Ling County, the task was more mandatory than voluntary, under the purview of the Administration of Industry and Commerce. From piloting to city-wide coverage of most, if not all entities in Ling, it took about one year from September 2021 to the third quarter of 2022, while it only took the first few months of 2022 in Beijing. This was partly determined by the power a government has, the level of urbanization (Sect. 4.5), and the nature of private participants. It goes without saying that the Beijing government has more power than the Ling government over its people and enterprises, and thus was able to implement the venue code more rapidly. Ling has many small, informal businesses that were not properly licensed and outside of administrative supervision, which increased barriers for mobilization.

Second, residents' willingness to comply was associated with their perceived risk of Covid exposure and their fear of enforcement by participating entities (often through roles like security guards and receptionists). This was why most Chinese cities were able to promote the tool more effectively in March and April 2022. The Shanghai outbreak provoked public fear and the central government's insistence on the zero-Covid objective channeled through the Tiao-Kuai hierarchy, compelled local governments and grassroot actors to commit in the rollout of the venue code and enforcement. With low mobility, Ling was able to control Covid cases well within the single digits until after August 2022, when the spread of the Omicron variants increased. As one of my Ling interviewees stated in October 2022: "*there are QR codes everywhere. It is convenient to scan. No troubles at all. The big data can follow everyone's whereabouts…The venue workers will remind you to scan the code but will not stare at you. You show them the health code after scanning the venue code. If your travel card looks normal, they will let you pass.*" When asked whether they felt uncomfortable scanning the codes, they said: "*I don't feel anything. These are all public venues.*"

Third, Chinese residents' acceptance of the venue code should be situated beyond the immediate disciplinary effects and perceived imperative to comply. Heightened dependence of Chinese people on platforms and apps such as WeChat and Alipay through a long-term habituation make them more likely to subject to digital enclosure (Andrejevic 2007; Byler 2021; Chen et al. 2018); these platforms and apps provide an effective digital infrastructure for government surveillance (Chen and Qiu 2019). It was through WeChat and Alipay that digital Covid surveillance could be rapidly enforced nation-wide. In addition to this context, many circumstantial factors such as waiting-time also mediate enforcement and compliance. Despite pervasive compliance, many residents viewed it as a de facto surveillance tool and were worried. These responses reflected a divided Chinese society regarding the zero Covid approach, shedding light on daily state-citizen negotiation and the nation-wide protests in late 2022.

4.5 Spatial (Re)organization

To facilitate the implementation of venue code, intense spatial (re)organization was deployed at the community level to define what a venue and an entrance is, changing and limiting urban mobility patterns and limiting access to urban resources. In this section, I will discuss ways how spatial organization was performed in various rural and urban settings and mediated public access to spaces and resources. Residential communities (*xiaoqu and/or shequ*) have become major spaces of government during post-reform China after rural communes and urban *danwei* were dismantled (Bray, 2005, 2009; Tomba 2014). Almost all its urban residents, over 900 million and about 64% of its total population, are living in gated communities (Hamama and Liu 2020). The spatial (re)organization of the University X in Beijing and Ling County People's Hospital exemplified that of many gated communities and public venues throughout China. Under the slogan of "one code for each venue and entrance" (Beijing Evening News 2022), in almost all Beijing communities, barricades at the entrances previously used to separate pedestrian and auto vehicles were extended further outward with checkpoints to allow rooms for social distancing for visitors waiting in queues while displaying their codes. These checkpoints were often facilitated with digital infrared thermometers, guards, and kiosks/stands for code reading. In order to reduce costs for hiring guards and facilitate easy management, very few gates, one or two at most, were open restricting mobility to central locations.

Additionally, from early 2022 with the emergence of campus cluster transmission, University X began to segregate its residential community from the campus with barricades, chained gates, metal panels, and barbed wire (Figs. 4.6, 4.7 and 4.8, taken in July 2022). Previously, these two spaces albeit walled were easily accessible from six gates from the outside and numerous passageways and roads from within without any access control. After the segregation, community residents could only enter and exit through the north-west and north gates, while re-entry was granted upon scanning the venue code and displaying a green health status. Meanwhile, students could enter and leave campus only through the east-south gate with permits rarely granted. Another three gates were locked and fenced. Only one passage between the two spaces was left for faculty and staff (not the students!) to enter the campus and perform their teaching and administrative tasks. Many public spaces doubled down and adopted draconian spatial reorganization measures. One extreme case occurred to a college in Zhengzhou City, Hebei Province, where authorities electrified the wire fences during a lockdown in September 2022, resulting a social media outcry among students complaining that the college treated them as prisoners (social media data verified by multiple sources). The college dismissed such complain and clarified that the signs with such a warning was for intimidation purposes and that the fences were not "really" electrified.

Compared to the gated communities, open, sprawling urban and rural spaces have significantly challenged the enforcement of the venue code among other measures. These spaces include but are not limited to rural and urban villages (Bach 2010), farmers markets, and special unwalled communities (i.e. Beijing *hutong*), which are

Fig. 4.6 The red-brick buildings were the residential buildings. This gate sealed with blue metal panels were once a crucial passageway between the campus and the residential community. On the right-hand side with green-color walls is a kindergarten (Xiaoling Chen, July 2022)

Fig. 4.7 On the left of the road was the campus, on the right was the residential community, and to the end of the road was the west gate of the university. The gate—used to have large traffic—was now locked. The orange sign on the barbed wire reads "prohibit from approaching and climbing" (*jinzhi kaojin, yanjin panpa*)—same as the yellow sign in Fig. 4.8. below (Xiaoling Chen, July 2022)

Fig. 4.8 This photo was taken from the residential community. Students had passed through the gaps between, under and over the gates and fences. Thus, the university placed barbed wire around them (Xiaoling Chen, July 2022)

spatially well-connected communities, spaces, and infrastructures. Spatial reorganization in such spaces can be illustrated by Guangzhou's urban villages during the outbreaks in October and November 2022 (Liang 2022) During the period, urban villages was walled by plastic barricades and temporary gates were erected, while houses and buildings inside were fenced off by plastic and metal panels confining residents to their homes. These spatial arrangements reduced the flexibility of human movements and interactions, directed human flows, and increased barriers to accessing urban resources distributed within and outside the community.

Spatial reorganization experienced constant changes over time, as state actors and residents negotiated, made compromises, reached agreements, and repeated the process as circumstances changed. Some gates were locked for a few days and reopened afterwards, leaving no physical traces, but causing much inconvenience and complaints. Most changes were remembered in various forms such as warning signs, barbed wire above and below the chained gates, and layers of different barriers. These all indicated attempts and failures on the part of their residents to break, to breach, or simply to protest. In November 2022, protests broke out in the urban villages of Guangzhou City where residents tore down barriers as a display of anger. When I left Ling and Beijing, more and more barricades and fences were being installed. Many cities had deployed more draconian and restrictive measures such as locking up residential buildings, and fencing off flats and apartments, to further prevent mobility and human interaction. Given these acts, we should think of spatial organization beyond its immediate functions and evaluate how it allowed local officials and grassroot entities to test the boundaries of their power, and public acceptance of intensifying restrictions.

4.6 Logistics and Technologies

> Teacher Lu paused at the passageway and pulled her mask to the chin. She fit her face to the facial recognition interface erecting on the top of one of the four turnstiles while putting her right palm against the automatic body temperature reader on top of the interface. After successful readings, the interface displayed her health code and body temperature and announced out loud her permitted pass. Teacher Lu went to the campus five days a week to conduct her teaching responsibilities to the de facto campus-confined students. Different from the other public venues in Beijing that require a 72-hour negative Covid result, this campus required a 48-hour result for access, frequenting Teacher Lu's routine tests during the weekdays.

The venue code directly and indirectly capitalized an array of increasingly integral logistical and technological infrastructure. To begin with, the practice was availed by linking the venue code system to extensive databases harvested from various surveillance systems (i.e. health code), Covid containment measures (i.e. mass testing and vaccination), and national epidemiological database (risk categorization of cities and communities based on the color-coded system). Paradoxically, the significance of the venue code to Covid containment on a daily basis was made prominent through updates of these databases on a *reasonable* frequency. Otherwise, the surveillance nature of the venue code would become very obvious. Put another way, strict enforcement of the venue code was legitimized by its positive association with Covid containment claimed by the Chinese government. In turn, continued identification of new cases necessitated the existence of such a practice, forming a mutually reinforcing cycle for the necessity of invasive surveillance tools and the zero-Covid policy. This was part of the reasons why many Chinese netizens believed that some outbreaks were "man-made crisis" (i.e. by manipulating positive case numbers and falsifying Covid test results) on the part of the local governments and private entities to legitimize excessive containment and to make profits.

Ling County and Beijing exemplified two common but different approaches towards mass Covid testing schemes. As a semi-urban, less affluent county, Ling only implemented mass testing when needed. Rounds of *voluntary* mass testing were documented in February and March 2022 when there were sporadic outbreaks with several cases in the region. *Mandatory* Covid tests were required during Fall 2022—local official media first announced that "avoiding Covid tests is a crime", indicating the heightened state control and crackdown on dissent against Covid test. In Beijing however, a 72-hour negative result was required for most of 2022 and more frequent tests (24-hour) were required at times. Originally many felt the need to get tested either because the tests were free and/or it was essential for the performance of their daily life functions, but later it became necessary to avoid legal liability. A tool that was deployed to prevent outbreaks and alleviate epidemic uncertainty, had itself become the source of risk for many residents, making residents the "first [if not only] responsible person" (*diyi zeren ren*) of their health.

Mass testing was normalized as a major Covid containment strategy from mid-2021 in cities such as Shenzhen and Nanjing. Repeatedly endorsed by the central government, it has since been propagated as the most effective containment effort at the lowest cost. According to Jinmin Li, the vice director of the Clinical Testing Center of National Health Commission, China's Covid testing capacity has increased from 1.26 million test-tubes per day in March 2020 to 57 million test-tubes per day in May 2022, involving an industry the size of ten billion *yuan*. Each test tube contains swab samples from one to ten people. Such an increase can largely be ascribed to the spread of Omicron variants and the detrimental containment failure in Shanghai between March and May 2022. Shanghai alone saw Covid test capacity increased from 1 million to 8.5 million test tubes per day within those months (Yu and Li 2022). Early May 2022 saw an eruptive expansion of Covid testing infrastructure, as Shenzhen's "15-min sampling circle" adopted in February 2022 was hailed as an effective way to curb Omicron spreads against the backdrop of the Shanghai *failure*; many cities followed suit even without any signs of an outbreak (Yu and Chen 2022). This "Great Leap Forward" of Covid testing (*hesuan dayuejin*) came to a halt in June as many cities found this measure difficult to sustain financially, when the central government forbid local governments from using residents' Health Insurance Funds or to charge them for mass testing. Some also realized the negative impacts of such normalization on the local economies as residents were more likely to stay at home and thus reduce economic activity (Yu and Chen 2022).

In addition, as various databases were shared and linked, the adoption of facial recognition and code reading technologies were possible in an all-in-one digital system (Chin and Lin 2017). All-in-one digital systems, generally called a "Covid Prevention Digital Sentinel" (*fangyi dianzi shaobing*), integrated, and combined the functions of visual capture, early warning, video coding, video recording, and/or data management. The adoption of these advanced technologies permeated public-funded venues such as universities, public transit systems, and public hospitals in large urban cities like Beijing, Shanghai and Guangzhou; in less financially viable places like Ling, manpower remained the major resource for the implementation of different surveillance projects. For example, University X installed digital sentinels in literally all residential, teaching, and administrative buildings, major passageways, and dining halls, purchased from the Xinkaipu Eletronics Co. Ltd.—a Chinese company listed on the Shenzhen Stock Exchange. Xinkaipu's machines were directly or indirectly linked to the health code database managed by local authorities (i.e. Guangdong, Beijing). Such a linkage allowed their machines to harvest a wide range of customized data points in real time by reading only one card/code/face. The company focused on three areas of business: smart campus, smart government and business, and social digitalization. During the first half of 2022, the company obtained 119 new clients for smart campus services due to the need for facial recognition, automatic body temperature reading, and code reading during a time of prevalent campus lockdowns (Xinkaipu 2022).

While digital sentinels with facial recognition were used in venues with a stationary population such as university campuses, code readers were more prevalent in public transits and hospitals. Beijing and Shanghai combined their health code, venue code, and travel pass (for busses and subways), achieving "access with one code" (*yima tongxing*). In Shanghai, this app was known as "Metropolis" (Xinmin 2022), while Beijing allowed linking a travel pass to its health app. Scanning the generated code on the code readers at a subway station allowed a resident to complete the requirements of displaying a green health code, pay for the travel fares, and have their location registered in real time. In this scenario, residents with a yellow- or red-coded health code or an expired Covid test result would be rejected access and "handled" by the site personnel according to the Covid prevention and control guidelines. Guards were placed to supervise entrances during operational hours, even when digital devices were installed that only granted access upon successful facial recognition and code reading. This was because they distrusted their targets (i.e. residents, staff, clients, and visitors) whom they suspected would go around or damage the machines if left unmonitored. This however, reflected the difficulty of getting public compliance with unpopular policies from excessive containment measures.

Digitalizing surveillance removed room for maneuvering on the part of residents, while improving information gathering capacity on the part of the local state. While there are many geographical variabilities in terms of how different codes were integrated and databases were shared and synchronized, technologies allowed the authorities to achieve real-name authentication, and to simplify the process of displaying/scanning various codes in the guise of citizens' health and safety. The "smartization" of Covid containment measures concealed both the presence and nature of different surveillance tools: harvesting in real time as many personal data points as possible, strengthening population control and discipline, increasing fear, and potentially manipulating data towards the interests of the state and local governments, which do not necessarily align with that of the people. This is exemplified by the most infamous case occurred in Spring 2022 in Henan Province, where local officials assigned red health codes to people travelling to Zhengzhou City to protest local banks freezing their accounts (Gan 2022).

4.7 Conclusion

This paper addresses implications of China's Covid containment regime and potential sites for state-citizen disputes through the lens of the venue code. While the previous sections have primarily discussed the impact of the venue code on Chinese people's everyday life, two cases reveal the insidious roles of the physical and technological infrastructures in mediating emergency management and state-citizen relationships. On July 09, 2022, a patient stabbed and injured four people in one of Shanghai's largest hospitals—Ruijin Hospital (Ng 2022). Social media pictures and a video clip illustrated the security guards' inaction, in the front of over ten health code reading machines obstructing the escape routes of dozens of fleeing patients, evoking heated discussion on the problematic installation and management of the machines. On November 24, a fire in Urumqi ignited unprecedented protests in China and abroad, as many believed physical barriers resulted from Covid containment hampered escape and delayed rescue vehicles. These discussions and protests also reflected genuine public empathy towards the dead and injured, and the public recognition of the significant impacts of restrictive containment measures. As the most recently implemented measure, the venue code embodied features of almost all extant containment measures (i.e. the organization structure and infrastructures) while having its own specificities, thus offering a vantage point to the impacts of the containment regime.

The venue code was built on digital platforms for Covid containment (i.e. health code) and beyond (i.e. WeChat and Alipay), and capitalized on well-developed big-data surveillance technologies such as facial recognition. The initial versions of the health code did not collect information on mobility, and the travel card only examined mobility at the scale of a city; however, the venue code scrutinized Chinese citizens at the scale of a venue in real time. In the past three years, local enforcers might shut down some community gates due to sporadic outbreaks, but they rarely segregated spaces at the granular level which the venue code allowed them to. My findings suggest that the venue code enabled local governments and grassroot entities to redefine what a venue and an entrance was, and thus helped increase residents' tolerance for more physical restrictions. Such an increase was further strengthened by popular narratives that ascribe outbreaks mainly to individuals, blamed for "going around [in an irresponsible manner]" (*daochu zou*), "having no self-awareness" (*bu zijue*), or "causing trouble for the neighborhood/the country" (*gei dajia/guojia tian mafan*). By the next half of 2022, as Chinese citizens were trained to bear the responsibility for limiting their own mobility, many were basically confined to the spaces of their home, work, and community. Because of that, many basic needs and services were delivered door-to-door or were provided in close proximity by community workers. Maintaining this level of containment has kept local governments, grassroot entities, and their residents under significant strain.

Like other Covid containment measures in China, the venue code was quickly abandoned at the end of 2022. Its residuals—rusted barricades and barbed wire littering the walls and gates—scarred communities, if not people's memories and lives. Despite its abandonment, the venue code—and other containment measures—deserves more intellectual engagement. This article draws attention to the implications of the venue code beyond its immediate functions and effects. The technical aspects of the tool—How was the data managed and processed? How was it different from other surveillance projects that collected time-location data? Was it effective in containing the pandemic?—remain unknown. We have yet to know how the large amount of data on human mobility patterns collected through the venue code might be used and to what end, raising many ethical questions of the digital surveillance. Additionally, digital surveillance's erosion of social trust is particularly concerning in low-trust societies like China (Stevens and Haines 2020; Wang 2020a, b; Zuboff 2019). In this sense, my study of the venue code can be viewed as a window into the broader and longer-term effects of big-data technologies on societies undergoing intense urbanization and social development, which go well beyond the immediate concern for public health.

Acknowledgements Funding for research was generously provided by the Society of Woman Geographers, and by the Department of Geography, Graduate School, and the Center to Advance Research and Teaching in the Social Sciences at the University of Colorado Boulder. I would like to thank three anonymous reviewers for their constructive comments.

References

Andrejevic M (2007) Surveillance in the digital enclosure. Commun Rev 10(4):295–317

Bach J (2010) "They come in peasants and leave citizens": urban villages and the making of Shenzhen China. Cult Anthropol 25(3):421–458

Bach J (2019) The red and the black: China's social credit experiment as a total test environment. Br J Sociol 71:489–502

Bartz D, Alper A (2022) U.S. bans new Huawei, ZTE equipment sales, citing national security risk. https://www.reuters.com/business/media-telecom/us-fcc-bans-equipment-sales-imports-zte-huawei-over-national-security-risk-2022-11-25/. Accessed 25 Dec 2022

Batke J, Ohlberg M (2020) State of surveillance. https://www.chinafile.com/state-surveillance-china. Accessed 10 Jul 2022

Beijing Evening News (2022) Beijing: one code for each venue and entry, guarantee accurate information for code scanner. https://weibo.com/1703371307/Ly960aoRh?type=repost. Accessed 10 Jul 2022

Berg B (2009) A Dramaturgical Look at Interviewing. In: Qualitative research methods for the social sciences. Allyn & Bacon, Boston, pp 101–157

Bray D (2005) Danwei space. In: Bray D (ed) Social space and governance in urban China: the danwei system from origins to reform. Stanford University Press, pp 123–156

Bray D (2009) Building 'community': new strategies of governance in urban China. In: Jeffreys E (ed) China's governmentalities. Routledge, pp 100–118

Byler D (2021) Terror capitalism: Uyghur dispossession and masculinity in a Chinese city. Duke University Press

Chen H, Greitens SC (2022) Information capacity and social order: The local politics of information integration in China. Governance 35(2):497–523

Chen X (2020) Spaces of care and resistance in China: public engagement during the COVID-19 outbreak. Eurasian Geogr Econ 61(4–5):435–447

Chen X, Oakes T (2023) Time-space companions: digital surveillance, social management, and abuse of power during the Covid-19 pandemic in China. Crit Asian Stud 1–24. https://doi.org/10.1080/14672715.2023.2191248

Chen Y, Mao Z, Qiu JL (2018) Super-sticky WeChat and Chinese society. Emerald Group Publishing

Chen Y, Qiu JL (2019) Digital utility: Datafication, regulation, labor, and DiDi's platformization of urban transport in China. Chin J Commun 12(3):274–289

Chin J, Lin L (2022) Surveillance state: inside China's quest to launch a New Era of Social Control. St. Martin's Press

Chu JY (2014) When infrastructures attack: The workings of disrepair in China. Am Ethnol 41(2):351–367

Datong Health Commission (2022) Authority response comes: Can the venue code replace the health code and Tonglu Tong? http://www.dt.gov.cn/dt12345/rxdt1/202204/654428a3ae9a4e67b4988ba1684988e9.shtml. Accessed 10 Jul 2022.

Erie M, Streinz T (2021) The Beijing effect: China's 'Digital Silk Road' as transnational data governance. N Y Univ J Int Law Politics 54:1. Available at SSRN: https://ssrn.com/abstract=3810256

Evans B (2023) China's government is buying Alibaba and Tencent shares that give the Communist Party special rights over certain business decisions, report says. Yahoo Finance. https://finance.yahoo.com/news/chinas-government-buying-alibaba-tencent-165617215.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAACbpKgqBiUSdqs8NbG9QWf4uG0t-Q18heIxPlwLfW3J9G8lIFMQua7THrv8lDDry-7bpbIidM2AMZnBkYXYOvkOcQxFcMbpy82mZLjjs7gHhcYmyCNTl5-NUVRmsGhi5oTtFLFTvEp41kEt20ipQlFqC3ghU-h_8ywVTfYSd-v30]. Accessed 29 Jan 2023.

Gan N (2022) China's bank run victims planned to protest. Then their Covid health codes turned red. https://www.cnn.com/2022/06/15/china/china-zhengzhou-bank-fraud-health-code-protest-intl-hnk/index.html. Accessed 15 Jul 2022

Gandy O (1993) The panoptic sort: a political economy of personal information. Boulder, CO: Westview Press

Gao Y, Cartier C (2022) The grid process: spatializing local governance in China. Eurasian Geogr Econ:1–26

Giddens A (1987) Social theory and modern sociology. Stanford University Press

Hamama B, Liu J (2020) What is beyond the edges? Gated communities and their role in China's desire for harmonious cities. City, Territory and Architecture 7(1):1–12

Herbert S (2000) For ethnography. Prog Hum Geogr 24(4)

Herzfeld M (1993) The social production of indifference: exploring the symbolic roots of Western Bureaucracy. University of Chicago Press, Chicago, IL

Hoffman S (2019) China's Tech-Enhanced Authoritarianism. Testimony before House Permanent Select Committee on Intelligence Hearing on "China's Digital Authoritarianism: Surveillance, Influence, and Political Control". https://www.congress.gov/116/meeting/house/109462/witnesses/HHRG-116-IG00-Wstate-HoffmanS-20190516.pdf. Accessed 10 Jul 2022.

Huang J, Tsai KS (2022) Securing authoritarian capitalism in the digital age: the political economy of surveillance in China. China J 88(1):2–28

Kharpal A (2022) China's next regulatory target—algorithms, the secret of many tech giants' success. https://www.cnbc.com/2022/01/07/china-to-regulate-tech-giants-algorithms-in-unprecedented-move.html. Accessed 29 Jan 2023
Lam T (2021) The people's algorithms: social credits and the rise of China's big (br)other. In: Mennicken A, Salais R (eds) The new politics of numbers: Utopia, evidence and democracy. London: Palgrave Macmillan, pp 71–95
Lees L (2004) Urban geography: discourse analysis and urban research. Prog Hum Geogr 28(1):101–107
Liang F (2020) COVID-19 and health code: how digital platforms tackle the pandemic in China. Soc Media + Soc 6(3):2056305120947657
Liang X (2022) Chinese migrant workers protest amid Covid-19 lockdowns in Guangzhou textile hub. https://www.scmp.com/news/china/politics/article/3199725/chinese-migrant-workers-protest-amid-covid-19-lockdowns-guangzhou-textile-hub. Accessed 16 November 2022
Lieberthal K, Oksenberg M (2020) Policy making in China. Princeton Univ, Press
Livingston J (2012) Improvising medicine: an African oncology ward in an emerging cancer epidemic. Duke University Press
Lukács G (2020) Invisibility by design: women and labor in Japan's digital economy. Duke University Press
Mann M (1984) The autonomous power of the state: its origins, mechanisms and results. Eur J Sociol/Arch Eur Sociol 25(2):185–213
Mittelstaedt JC (2021) The grid management system in contemporary China: grass-roots governance in social surveillance and service provision. China Information
Neuendorf KA (2017) The content analysis guidebook. Sage
Ng T (2022) Four injured in knife attack at Shanghai hospital. https://www.scmp.com/news/china/politics/article/3184731/four-injured-knife-attack-115-year-old-shanghai-hospital. Accessed 10 Jul 2022
Qiu JL (2016) Social media on the picket line. Media Cult Soc 38(4):619–633
Qiu JL, Bu W (2013) China ICT studies: a review of the field, 1989–2012. China Rev 13(2):123–152
Stevens H, Haines MB (2020) Tracetogether: Pandemic Response, Democracy, and Technology. East Asian Sci, Technol Soc: Int J 14(3):523–532
Su Z, Xu X, Cao X (2021) What explains popular support for government monitoring in China? J Inf Technol Polit
Suzhou Daily (2022) "The venue code" secures pandemic prevention. https://www.suzhou.gov.cn/szsrmzf/szyw/202204/fe697358802143a29716b3e659178ea8.shtml. Accessed 10 Jul 2022.
Tang B (2018) Deliberation and governance in Chinese middle-class neighborhoods. Jpn J Polit Sci 19(4):663–677
Tang B (2020) Grid governance in China's urban middle-class neighborhoods. China Q 241:43–61
Tarrow S (2018) Mann, war, and cyberspace: dualities of infrastructural power in America. Theory Soc 47(1):61–85
Thacker SB, Berkelman RL (1988) Public health surveillance in the United States. Epidemiol Rev 10(1):164–190
The Communist Party's Central Committee (CPCC), The State Council (2021) Notification of the State Council to Strengthen Grassroot Governance Mechanisms and Their Modernization of Governance Capacity. http://www.gov.cn/zhengce/2021-07/11/content_5624201.htm. Accessed 10 Jul 2022
Tomba L (2014) The government next door: neighborhood politics in urban China. Cornell University Press, Ithaca, NY
Van der Geest S, Finkler K (2004) Hospital ethnography: introduction. Soc Sci Med 59(10):1995–2001
Wang E (2023) 回望2022: 我們問了7位中國人, 去年掙了多少?過得還好嗎? https://theinitium.com/article/20230127-mainland-life-for-the-common-people-2022-review/?utm_source=sendinblue&utm_campaign=basic0129&utm_medium=email. Accessed 28 Jan 2023.

Wang F, Yin H, Zhou Z (2012) The adoption of bottom-up governance in China's homeowner associations. Manag Organ Rev 8(3):559–583
Wang X (2020a) Blockchain Chicken Farm: And Other Stories of Tech in China's Countryside. Farrar, Straus and Giroux Originals x Logic
Wang Y (2020b) In China, the 'Great Firewall' Is Changing a Generation. Politico. https://www.politico.com/news/magazine/2020b/09/01/china-great-firewall-generation-405385. Accessed October 10 2022
Warf B (2011) Geographies of global Internet censorship. GeoJournal 76:1–23
Wen Y (2021) Branding and legitimation: China's party diplomacy amid the COVID-19 pandemic. China Rev 21(1):55–89
Xinkaipu (2022) Financial Report January-June 2022. Via Finance Ifeng: https://app.finance.ifeng.com/data/stock/ggzw.php?id=714762523992&symbol=300248&ggtype=20. Accessed 24 Oct 2022
Xinmin (2022) Shanghai subway piloted the Metropolis App today, one screen displaying Covid test + health code + trave fare. http://news.xinmin.cn/2022/05/30/32175717.html. Accessed 24 Oct 2022
Xu B (2017) The Politics of Compassion. Stanford University Press, In The Politics of Compassion
Yan S (2019) The village testing China's social credit system: driven by big data, its citizens earn a star rating. https://scmp.com/magazines/post-magazine/long-reads/article/3012574/village-testing-chinas-social-credit-system. Accessed 10 Jul 2022.
Yang G (2021) Online lockdown diaries as endurance art. Ai & Society, pp1–10
Yang G, Wang W (eds) (2021) Engaging social media in China: platforms, publics, and production. MSU Press
Ye Y (2021) To improve governance, China Ups control over its grassroots. https://www.sixthtone.com/news/1007977/to-improve-governance%2C-china-ups-control-over-its-grassroots. Accessed 10 Jul 2022
Yu H, Li S (2022) The great leap forward of tens millions of Covid testing industry: no future after getting rich in one night. https://mp.weixin.qq.com/s/oEM5wCM48rSBUyUgYIIM2w. Accessed 26 May 2022
Yu H, Chen X (2022) Nucleic acid "Great Leap Forward" reduced temperature, multiple areas halt normalization of mass testing, what does it imply? https://mp.weixin.qq.com/s/qvVWek1jazhcahCxwmoRqg Accessed 23 June 2022
Yuan L (2022) Bumingbai Podcasts with Guoguang Wu, Xia Cai, Dongyan Jin, Chenggang Xu, Xianfeng Yin, Ying Yang. https://www.bumingbai.net/. Accessed on different dates in September and October 2022
Zhang L (2010) In search of paradise: middle-class living in a Chinese Metropolis. Connell University Press, Ithaca, NY
Zhang X (2022) Decoding China's COVID-19 health code apps: the legal challenges. In Healthcare 10(8):1479. MDPI
Zhou B, Zhong Y (2021) Social media, collective mourning and extended affective space: a computational communication analysis of Li Wenliang's Weibo comments (2020–2021). Int Mass Media 43(3):79–106
Zhou SL et al (2021) Lessons on mobile apps for COVID-19 from China. J Saf Sci Resil 2(2):40–49
Zuboff S (2019) The age of surveillance capitalism: the fight for a human future at the new frontier of power. Hatchette, New York

Chapter 5
The Platformization of Public Participation: Considerations for Urban Planners Navigating New Engagement Tools

Pamela Robinson and Peter Johnson

Abstract Professional urban planners have an ethical obligation to work in the public interest. Public input and critique gathered at public meetings and other channels are used to inform planning recommendations to elected officials. Pre-pandemic, the planning profession worked with digital tools, but in-person meetings were the dominant form of public participation. The pandemic imposed a shift to digital channels and tools, with the result that planners' use of technology risks unitizing public participation. As the use of new platforms for public participation expands, we argue it has the potential to fundamentally change participation, a process we call *platformization*. We frame this as a subset of the broader emergence of platform urbanism. This chapter evaluates six public participation platforms, identifying how the tools they provide map onto key participation frameworks from Arnstein (1969), Fung (2006), and IAP2 (2018). Through this analysis, we examine how the platformization of public participation poses ethical and scholarly challenges to the work of professional planners.

Keywords Urban planning · Platform urbanism · Public participation · Smart cities

5.1 Background

The shift to digital forms of communication in society has only accelerated as a result of the COVID-19 pandemic. Across many dimensions of public and private life, we have turned to digital tools to maintain and develop what were once largely in-person networks and social groupings. This shift is mirrored in how professional planners connect and communicate to residents. While the use of digital tools and

P. Robinson (✉)
School of Urban and Regional Planning, Toronto Metropolitan University, Toronto, Canada
e-mail: pamela.robinson@torontomu.ca

P. Johnson
Geography and Environmental Management, University of Waterloo, Waterloo, Canada
e-mail: peter.johnson@uwaterloo.ca

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_5

platforms undoubtedly removes some barriers to participation, it simultaneously inserts new ones. This raises questions about where digital platforms fit according to dominant planning frameworks of participation. To address these questions, we present an environmental scan of how large Canadian cities are currently, in 2023, delivering urban planning participation and engagement efforts vis-à-vis technology to their residents. We reflect upon how the use of digital technologies can unitize resident-planner interactions, exploring the impacts of public participation and civic engagement technology platforms for planners. This chapter concludes with future research and ethical considerations for planning practitioners and researchers.

5.1.1 Urban Technology Platforms

From Uber to Airbnb to SkiptheDishes, it's hard to live in a city and not bump into platform-based businesses. A platform business is defined as "bring[ing] together individuals and organizations so they can innovate or interact in ways otherwise not possible" (Kennedy, 2015 as cited in Knee 2021). While the platform businesses pre-date the internet, their rise is easily tracked since 2000 (Knee, 2021). The rise of technology platforms is driven, in part, by private sector interests in the data they gather (Lee et al. 2020). In urban settings, technology platforms have a significant impact on city pattern, form, and governance. Platform urbanism is defined as the "study and design of urban practices that respond to new configurations between physical and digital layers of people, networks, and urban infrastructure, resulting from real-time, ubiquitous technology and platforms" (Barns 2015, p. 7). Urban digital platforms such as Uber and Airbnb have entrenched themselves in cities worldwide (Leszczynski 2020; Rosenblat 2018). These platform companies have emerged as intermediaries that shape our urban behaviors, ranging from the way we travel or book accommodation (Gurran and Phibbs 2017; Barns 2014), to directly impacting the process of urban governance (Barns 2020; Gorwa 2019; Van Doorn 2019).

As professionals, urban planners are ethically obligated to uphold the public interest, need to actively consider how these platforms impact the privatization of municipal services and serve as a new form of public–private partnership (AICP, 2021; CIP 2016). Platform urbanism companies control proprietary algorithms and data sets that inform service optimization outcomes (van der Graaf and Ballon 2019; Mansell 2015). Many early studies of algorithmic bias demonstrate disproportionately negative outcomes for people from poor, diverse, historically marginalized and equity-deserving communities resulting from decisions taken based on results from these tools (Brossard 2018; D'Ignazio and Klein 2020; Eubanks 2018; McIlwain 2020). Critics raise concerns that these platforms accelerate and expand new opportunities for private sector firms to access municipal decision-making, through the implementation of a shared urban governance (see Aguilera et al. 2019). In this regard, concerns about platform urbanism are similar to more widespread concerns about smart city technology adoption and reimagining of the "city as platform" which critics claim act as an accelerant for the privatization of public services (Goldsmith and Kleiman 2017; Marvin and Luque-Ayala 2017; Leszczynski 2020). This concern is persistent, and the encroachment is challenging for local governments to resist insomuch that these platforms are products of global technology firms with robust government relations and lobbying budgets. As governments face significant social, economic, and ecological challenges and constraints in a post-COVID environment, the potential and promise for new forms of private sector partnerships bringing funding will be tempting.

Platform urbanism tools present new opportunities for government. The data platforms collect can offer governments the ability to use analytical tools to uncover new local insights that support data-driven decision-making (Gessa et al. 2019), and potentially enhance two-way communication between citizens and government, increasing openness, access, and transparency (Gagliardi et al. 2017; Janssen 2012). There are also examples of increased data collection through platforms leading to improved and more efficient service delivery and more informed public policy-making (Höchtl et al. 2016). However, the potential of platforms captured in academic literature seems to oscillate between optimism and fear that they will produce "centralized, hierarchical system of urban control (Barns 2020 p 22)". Research has raised critical concerns around data privacy and security, and the potential for mass surveillance and greater corporate influence (Barns 2020; Kitchin and Dodge 2019). Beyond the collection and ownership of data, there are potential challenges with the quality of data that underlies platforms, as it may be biased, non-representative, or inaccurate (Green 2019; Matheus et al. 2018). The same is true of algorithms, through which decision-making happens within an impenetrable black box, leading to a lack of accountability and transparency with respect to outcomes (Kemper and Kolkman 2018). Additionally, there are concerns around the proprietary nature of the algorithms behind the platforms and the opaqueness of their systems (Fields et al. 2020). Research suggests that this may lead to the amplification of existing inequalities experienced by underserved communities, due to a higher likelihood of those communities not being appropriately represented in the data or adequately considered in the creation of algorithms (Dickinson et al. 2019; Wiig 2015). Many of these risks stem from platform firms' focus on delivering efficiency, when in fact, some of the most important work of government does not lend itself to efficiency because addressing challenges including inequity, sustainability, justice and democratic engagement take time and require relationship building (Corbett et al. 2019; Green 2020). Given this tension, the ways in which urban platform tools privilege data quantity over quality require further attention.

This chapter provides an initial overview of the increased platformization of public participation in planning, as a direct consequence of the COVID-19 pandemic shift towards the use of online public participation tools (Robinson and Johnson 2021). We identify six key public consultation and engagement convenor platforms and describe the types of tools that these platforms provide for use by professional planners. We map these tools to three frameworks that evaluate opportunities for resident participation by Arnstein (1969), Fung (2006), and IAP2 (2018). Through mapping these tools via frameworks of participation we show the broad positioning of platform tools towards more limited, unitized forms of participation. The chapter concludes with reflections on how the platformization of participation has research and practical implications for urban planners.

5.1.2 Public Participation in Planning

Urban planners have wrestled with how to conceptualize the participation of members of the public in the planning process for over 50 years. For example, the Arnstein (1969) Ladder of Participation serves as an enduring reminder that the mechanics of the participation process can be used for a full range of intent, from democratic to manipulative. Fung's Democracy Cube (2006) is a three-axis tool for analyzing myriad participation opportunities in response to complex governance challenges. The IAP2 Spectrum of Participation (2018) breaks down how organizers can reach different depths of engagement with community members using a range of tools and techniques. Scholarly planners conduct research on the conceptual and practical aspects of public participation, and this guidance is important for professional planners whose ethic of working in the public interest is enshrined in the American Institute of Certified Planners Code of Ethics and Professional Conduct (2021) and the Canadian Institute of Planners Code of Professional Conduct (2016). Yet, despite these deep ethical, theoretical, and applied commitments to improving approaches to public participation, the relationship between planners and residents is one that continues to generate durable challenges.

The ways in which planners have positioned themselves in relation to residents are evolving. Professional planners have moved from holding a perceived monopoly on expertise towards acting as facilitators, and now an emerging professional role as active participants in undoing systemic patterns of colonialism and racism (Goetz et al. 2020). Furthermore, practitioners and researchers continue to develop and implement new techniques and processes for participation and engagement. Public participation GIS, citizen science, volunteered geographic information, citizen juries, citizen labs, and world cafés are all examples of new ways to bring community members' ideas into a formalized planning decision-making process (Horgan and Dimitrijević 2019). In parallel, over the last 20 years, the turn towards digital communications has greatly impacted how planners consult and engage (Calzada 2018; Cardullo and Kitchin, 2019). Municipal governments are widespread users of social media platforms like Facebook and Twitter (Robinson and DeRuyter 2016). There is growth in private sector digital engagement platforms such as Bang the Table, MetroQuest, PlaceSpeak and SeeClickFix (Kleinhans et al. 2015). Proponents of these tools pitch their abilities to reach more people, more conveniently, thereby proposing to strengthen consultation and engagement efforts (Gagliardi et al. 2017). Critics flag concerns about digital inequalities excluding community members and technology's capacity to unitize and quantify resident contributions, thus masking its potential heterogeneity, impact, and value (Silva et al. 2019; Johnson et al. 2020). Streich's pre-pandemic predictions (2018) regarding how new technology tools will create new opportunities and challenges for planners. One running theme through his ten theses is his call to planners to actively consider how changes in technology and information flow will require our profession to adapt and change, moving to less hierarchical and more bottom-up engagement by our residents. Particularly resonant is his suggestion that planners be open to our community members playing a stronger leadership role in the generation of new planning ideas including sharing this thinking through technology platforms that enable public participation, hereafter called convenor platforms.

5.1.3 Convenor Platforms as Mediators of Public Participation

Given the pandemic-driven shift towards online participation, it is important to consider how resident participation in government decision-making is shaped by the tools used to facilitate those conversations. Convenor platforms can be used to enable interactions, but their design and the tools they offer channel the participation process and its substance in particular ways (Amoore 2016). This tension is revealed in smart city projects, where participation can often be limited to data inputs such as likes or voting ideas up or down into opaque algorithms or decision-support systems (Sieber et al. 2016; Pasquale 2015).

This research questions whether, with the use of certain platforms, specific limits are placed around the types of contributions that participants can make. For example, if the social media site Twitter were to be used within a consultation process, are there technical limits to what can be shared, a possible lack of geographical boundaries around who is involved (due to the global nature of Twitter), and data quality concerns based on anonymity of contributors? Similar concerns can be raised through the use of many types of distributed online participation systems, where public participation goals are restricted or formed based on the structure and nature of the technology itself. One concern raised by Johnson et al. (2020) is that through the use of participation technologies, contributions become more unitized and transactional. Transactions could be considered as those that are harvested passively from embedded sensors, for example collecting data on foot traffic, and also those that are intentionally made by communities within a structured planning process. The type of tool used frames the mode of participation; for example, allowing participants to upvote options or provide limited text-based comments to proposals online are unidirectional transactions, without providing participants the opportunity to engage in dialogue or propose alternatives. While this type of data may be useful for information or basic consultation, it stands in contrast to more deliberative methods of gathering contributions that may be better able to achieve goals of public participation identified by IAP2 (2018), such as involvement, collaboration, or empowerment. Planners must choose their technology tools carefully to ensure that the functions provided align with the public participation goals.

5.2 The Rise of Participatory Convenor Platforms

To support the process of public deliberation around government decision-making, planners are increasingly making use of *convenor platforms*, which we define as a platform urbanism tool that serves as a service and access intermediary between governments and their residents. In Table 5.1, we highlight six widely used convenor platforms and include excerpts from their marketing pitches as a way to contextualize their services within the frame of public participation. These platforms are reported to be in use in Canada, the United States, Australia, and in some European cities:

- ***The Hive:*** https://the-hive.com.au/: "Instantly reach more people, and make more informed decisions today."
- ***Bang the Table/EngagementHQ*** (under Granicus umbrella): https://go.engagementhq.com/ "One robust platform lets you inform, engage, measure and build community through meaningful relationships and on-going interactions."
- ***PlaceSpeak***: https://www.placespeak.com "By participating in public consultations and building stronger neighborhoods, PlaceSpeak empowers you to make a meaningful impact on the communities where you live, work, and play."
- ***CitizenLab***: https://www.citizenlab.co/ "Uncover community insights to make inclusive and data-driven decisions in your city. CitizenLab's centralized community engagement platform makes it easy for governments to engage their residents, manage input, and make informed decisions."
- ***Nextdoor***: www.ca.nextdoor.com "By bringing neighbours and organizations together, we can cultivate a kinder world where everyone has a neighbourhood they can rely on." "The Nextdoor for Public Agencies program allows various government entities including cities, counties, police departments, and fire departments to launch Nextdoor neighborhoods across their municipality."
- ***Decidim***: https://decidim.org/ "Decidim helps citizens, organizations and public institutions self-organize democratically at every scale." "Decidim is a platform for citizen participation made by the people and for people. Its source code is open and can be inspected, modified, and enhanced by anyone. The Decidim software is covered by the *AGPL license*."

All of the tools, with the exception of Decidim, are third-party vendor tools, designed to be purchased via contract by a public agency. Decidim, which in Catalan means "let's decide", is an open-source digital decision-making tool (Barandiaran et al. 2018), allowing municipal staff, after agreeing to the platform's social contract, to deploy it without needing to procure access or professional services from the platform provider. Of the six tools, Nextdoor may appear to be an outlier because it is more commonly known as a neighbor-to-neighbor social network; however, this technology company also offers a version geared towards public agencies, pitching the tool as an effective means of "reaching the right audience" with geo-targeted messaging. All of these tools are designed to be used by public officials to provide a platform for a range of public consultation and engagement activities. Implicitly and explicitly, these tools are being sold on the promise that they can increase the

Table 5.1 Comparison of Convenor Platforms

Frameworks for Participation			Tool functions	Convenor Platforms					
Fung (2006)	Arnstein (1969)	IAP2 spectrum (2018)		Hive	EngagementHQ	Citizen Lab	Nextdoor	Decidim	PlaceSpeak
Communicative influence	Informing	Inform	messaging			✓	✓		
			information posting/blogs/pages			✓	✓	✓	✓
			official updates			✓		✓	
			storytelling	✓	✓			✓	
			Q&A	✓	✓				
			discussion board	✓			✓		✓
			forum	✓	✓			✓	✓
			competitions/voting ideas up or down	✓				✓	✓
			geotargeted messages	✓			✓		
			resource library					✓	✓
			event calendar					✓	✓
			project timelines/process tracking					✓	✓
			process map for participation						✓
Advise and consult	Consultation	Consult	survey	✓	✓	✓		✓	✓
			poll	✓	✓	✓		✓	✓
			petitions		✓				
			guestbook with moderation/debates		✓			✓	✓
Co-governance	Placation	Involve	option/scenario analysis			✓			

(continued)

Table 5.1 (continued)

Frameworks for Participation			Tool functions	Convenor Platforms					
Fung (2006)	Arnstein (1969)	IAP2 spectrum (2018)		Hive	EngagementHQ	Citizen Lab	Nextdoor	Decidim	PlaceSpeak
			social mapping/place based commenting	✓	✓	✓	✓		✓
			online workshops/assemblies/conferences			✓		✓	
	Partnership	Collaborate	brainstorming/ideation/crowd sourcing/post-it notes	✓	✓	✓		✓	✓
			citizen proposals			✓		✓	
Direct authority	Delegation	Empower	participatory budgeting	✓		✓		✓	
	Citizen control		change the tool/add functionality					✓	

Note Convenor platforms change their tool functions over time. This mapping of functions was completed in January 2023 number of participants, expand the opportunities for these residents to contribute, and ultimately strengthen democratic decision-making through their use.

Table 5.1 provides an overview of the variety of tools these platforms offer to planners for working with their community members. There are similarities and differences in the tools and functionalities. For over 50 years, planners have been challenged and criticized for hosting public events that are more performative than democratic, giving the illusion that resident participation might actually make a difference. From the Ladder of Participation (Arnstein 1969) to the Democracy Cube (Fung 2006) to the Spectrum of Participation (IAP2 2018), critics of planning processes have evaluated the extent to which different participation and engagement techniques can lead to more or less resident impact on planning decision-making outcomes differentiating between them. These three taxonomies all share a similar approach, assessing information provision as the lowest form of planner-resident interaction with consultation and involvement in the middle, and ranking partnership, collaboration, empowerment, citizen delegation, and/or control of decision-making as the highest. The CitizenLab tool provides a framework by which they compare their platform's functions against the IAP2 Spectrum of Participation elements. In Table 5.1 we map all of the functions of all six convenor platforms across the Arnstein, Fung, and IAP2 frameworks. This mapping exercise is instructive as it helps inform the evaluation of the potential impact of these tools on the nature and impact of public participation.

Our first observation is that each platform supplies tools for information provision functions, including blogs, discussions boards, Q&A functions, official notifications, direct messaging, project timeline trackers, and participation process maps that delineate how community members can participate. Next, these platforms, in all instances, support consultation activities through tools such as surveys, polls, petitions, and moderated discussion forums. These kinds of information provision and consulting activities are the basics of planner-resident interactions. For planners, having access to a platform that gathers these functions together in a user-friendly interface has clear advantages. For residents, a one-stop visit for a project is also easier to navigate than standard municipal sites sharing information about public meetings.

As we move to further explore platforms that offer tools to enable more interactivity, opportunity, and transfer of power to residents, we start to see more variation and fewer options. Hive, EngagementHQ, CitizenLab, Nextdoor, and PlaceSpeak all have tools that let users get more involved in their communities through participating in social mapping. This connection between input and geolocation creates a more impactful opportunity for residents to contribute but it is not as transformative as collaborative or co-designed engagement activities. A few tools include options/scenario analysis evaluation, citizen assemblies, and conferences. When considering how the platforms support relationships between residents and planners that are more collaborative or partnership-focused, we see fewer and fewer options. These kinds of resident-planner relationships are mobilized through activities such as crowdsourcing and ideation activities. Both CitizenLab and Decidim have functions that allow residents to propose their own solutions.

There are only three platforms (Hive, CitizenLab, Decidim) that offer a tool to support participatory budgeting, which is a process that has the potential to reach the empower/delegation/direct authority level of planner-resident interaction (Karner et al. 2019). And, most notably, Decidim is the only tool, through its design, license, and structure, that potentially allows residents to evolve the tool through creating new modules. This open-source model allows Decidim to more closely align with the principles of open government, that is, that tools and processes of government be open to inspection by all impacted by the decisions made (Barandiaran et al. 2018). In this way, rather than requiring the corporate platform owner to develop or adopt any unique user needs or context, the open-source nature of Decidim allows for potential customization of a local implementation, drawing on premade modules from other governments or implementations, or through true custom development of a new module. Decidim also has the unique advantage of a global community of users, providing an opportunity for the sharing of best practices, lessons learned, and other implementation challenges. For a government that chooses to use Decidim, there is less reliance on an external, profit-motivated vendor, allowing more technical and operational expertise to remain in-house, increasing the level of control that a government may be able to exert over data collection (Johnson and Scassa 2023). This is particularly relevant as governments develop and refine their own data governance policies that may require specific data protection and retention arrangements.

An examination of the functionality included in these six platforms reveals a range of opportunities provided to planners for information provision and consultation activities. For both planners and residents, there are clear practical benefits to gathering these functions in one platform that is both easier to procure and simple to deploy. These platforms sell themselves as helping planners convene processes that are more impactful, more democratic, and will potentially achieve better decisions. But when we map the tool functions against the Arnstein, IAP2, and Fung engagement frameworks, we see the majority of the platforms' tools focus on information, consultation, and involvement forms of participation, all areas where communities have less power and influence over decision-making. This reflects the unitization of participation as described by Johnson et al. (2020), where technologies employed place limits around the type of participation. Though there are some examples amongst the six platforms where tools are provided to support opportunities in which residents have more power and agency, there are certainly fewer provided. This shows that there is still much work to be done in the use of platform technologies to support these more complex aspects of public participation, including foundational questions about the appropriateness of technology for these tasks.

5.3 The Platformization of Public Participation: Questions for Planning Practice and Research

The evaluation of these six convenor platforms leads us to question how the platformization of participation could impact the ways in which planners and residents interact, and the role of technology as a conduit for that participation. Within this shift towards technology-mediated participation, there are clear implications for planning research and practice. When considering the relationship between the planning profession and technology adoption, one striking observation about these platforms is that their sales pitches (Table 5.1) are similar to those of other smart city technology vendors who offer efficiency and integration through dashboards and analytics (Robinson and Biggar 2021). Another similarity is that the majority of these tools are proprietary software, developed and supported by third-party, private sector vendors. The exception here is Decidim, which is an open-source, free tool developed by the City of Barcelona. Both approaches (proprietary vs. open source) to participation platforms have challenges and benefits, as the adoption of enterprise technology within government is a complicated process, impacted by organizational as well as technological factors (Janssen et al. 2012; Khan and Johnson 2020). Given this, there are potential cues from the adoption of other technologies within planning that can guide government directions on platforms specifically, and with platform urbanism more generally.

The challenges to participation technology adoption within planning can be seen in earlier research on how smart city platforms position residents vis-à-vis city hall, raising concerns about the potential for the technology to make these interactions more transactive than transformative (Calzada 2018; Johnson et al. 2020). Participation platform tools can potentially make it easier for municipal staff to reach residents, but these tools do not necessarily generate better-quality participation or deeper levels of engagement. Previous research has concluded that in-person consultation leads to the creation of richer ideas, and when it comes to online participation, quality of contributions tends to be low (Einstein et al. 2022). This same literature states that digital engagement does not necessarily expand the scope of the conversation, and reaches the same educated, well-informed residents that would typically attend an in-person consultation (Pina et al. 2017). While digital consultation processes are potentially more accessible to those with access to technology, this medium strips away opportunities for critical debate and meaningful discussion between individuals with different life experiences which are fundamental parts of collaborative planning theory and practice. As noted earlier, the inequities and challenges present with in-person participation persist, and may be further exacerbated by digitally enabled processes. For example, with their internal dashboards that track resident interactions, platforms make it easier to focus on quantitative evaluations of participation than the qualitative richness (or not) of the contributions. Given the complexity of many urban planning projects, planners need to strategically select their technology tools to ensure the tools allow for the depth of resident deliberation needed for the planning decision. Planning research and practice needs to be more active collecting data and analyzing how the expanded use of digital convenor platforms impacts the depth and breadth of resident participation.

In light of these comparisons, planners seeking to work with these tools should note research about broader trends in platform urbanism and lessons learned from early smart city experiments (Robinson and Biggar 2021). A number of important questions emerge. First, whose interests are served when platforms are used and who benefits from the platformization of participation? These tools purport to make it easier for the platform host to do their work (e.g. the municipal planner), but what do we lose when we ask the public to adapt to our process by participating via convenor platforms? When we migrate participation onto platforms, what happens to dissent? How is dissent registered when the platforms channel participation in particular ways? It is important to explore how dissenting members of the public register their viewpoints (Legacy 2017). Of the six platforms reviewed here, CitizenLab and Decidim have functions that allow residents to submit proposals, and Decidim is the only one that allows participants to build new modules to change the platform and the process it supports itself. If the platform itself doesn't support dissent or digression, how do dissenting points of view find their way into the public deliberation process? If dissenting members of the public migrate to other platforms to organize, then this input may not necessarily find its way into the public process (e.g. see Evans Cowley, 2010). Members of the public have many platforms at their disposal, but there is no guarantee that these other contributions will be recorded or accepted in the formal public process via the platform of record. Future research should track how and where dissent is registered and considered when digital convenor platforms are implemented.

When hosting a public conversation, planners must account for who was consulted or engaged with the expectation that more input and contributions are better and more democratic. This accounting includes both official channels (such as surveys) and unofficial channels, such as social media. There are occasions in practice when planners find themselves on the receiving end of significant volumes of public input. Convenor platforms have the potential to further add to this volume of participation. Scholars have long been concerned about the self-reinforcing nature of technology (see Ellul 1964). Convenor platform use by planners may necessitate the need for automated processing of public input because of the ways in which these platforms and their tools make it easier for planners to inform and consult their residents, resulting in higher volumes of contributions. Natural language processing (NLP) technology is now being used by some government staff to process and evaluate these public contributions (e.g. see Government of Canada 2020). As participation is sought via convenor platforms offering efficiency, the extent to which NLP is used behind the technology interface is not easy to discern. The use of this technology to evaluate public contributions is an area of future research.

Planners need to carefully consider how automation impacts their evaluation of public input. The potential for bias in algorithmic programming of artificial intelligence tools is well documented (see Brossard 2018; Buolamwini and Gebru 2018; O'Neil 2016, among many). Municipal staff need to actively consider how bias may impact both the quality of the public input and how the use of these tools impacts public trust in the process itself (eg. see Benjamin, 2019). The efficiency and analytic opportunities that NLP offers planners raises further questions. When public contributions arrive via a third-party platform, are government staff aware of whether the platform is using technology tools including NLP? To what extent can government staff do their due diligence with regard to bias in the tool when the tool itself is being provided by a third party? Are these third-party vendors transparently disclosing the uses of these kinds of technology? And are they open to discussions about potential bias?

Robinson (2022) asked: Is there a civic obligation to disclose when public input is being filtered through automation? This question was raised based on the assumption that when community members participate in a government-led public participation or engagement process, a politician or staff person will be receiving and reviewing their inputs. Late in 2022 when ChatGPT was released, agencies began to experiment with bot-generated content and the ethical obligation to transparently communicate when technology instead of humans did the writing. Would our residents participate in the same way if they knew their feedback was going to a machine instead of a person? The proposed efficiency of these technology tools will be tempting to cash-strapped and understaffed government offices consulting and engaging their publics. But the intermediary roles this technology plays needs further attention and vigilance if the civic intent of the consultation and engagement processes is to be upheld. Planners have called upon our professional organizations to consider revisions and updates needed from technology advances (e.g. see Schweitzer and Afzalan, 2017; Robinson 2022). But as of now, the current codes of ethics for both Canadian and American professional planners are silent on how planners should consider new technology use in the context of their work in the public interest.

5.4 Conclusions

This chapter investigates six convenor platforms that are used as conduits for public participation. We track how the specific tools offered within each platform meet the goals of public participation as theorized within the Arnstein (1969), Fung (2006), and IAP2 (2018) participation frameworks. The majority of tools offered via these six platforms fit the informing and consulting roles of planning, yet largely do not address the engagement and power sharing roles identified in the participatory framework literature. We make that argument that urban planners should take time to reflect on the nature and impact of new digital trends in public consultation and engagement.

Practicing planners and researchers know that despite the proclamations about the benefits of consultation and engagement technologies, there are also problems and challenges with their use in public consultation and engagement efforts. Rather than default to the assumption that technology makes it easier for residents to participate, urban planners have an ethical and practical obligation to think about how the public interest is kept when these tools are used. We need to pay attention to how these platforms are designed and function so that the use of the technology does not work at cross purposes to the inclusion and democratic goals of resident participation in the planning process.

Whether working inside city hall guiding municipal consultation and engagement efforts or working in private practice as consultants, all registered/certified professional planners are bound by the same ethical codes of practice. Currently, there is a paucity of guidance in professional codes about the ethical obligations for planners when working with technology and data. There are several initiatives underway that seek to remedy this situation. The Cities for Digital Rights movement is making great strides towards helping local governments work more strategically and ethically with technology. Its member cities are also doing important work leading with privacy frameworks and their shared efforts to reduce the need for cities to duplicate their approaches to data privacy. Despite these advances, data, privacy, and rights initiatives are commonly held in the chief information or technology offices of local governments, which makes us consider how and if professional planners are engaged in these developments. It is important to remember that urban planners are at the institutional intersection of public consultation and engagement technology and how it is used to mediate conversations with residents. To fully uphold our ethical obligations, urban planners in practice and research need to more robustly engage in critical conversations about how convenor platforms impact resident-planner conversations quantitatively and, more importantly, qualitatively. Moving forward, working in the public interest means rolling up sleeves when it comes to addressing the platformization of participation, and the challenges that the adoption of this technology can generate.

Acknowledgements This book chapter draws on research supported by the Social Sciences and Humanities Research Council.

References

American Institute of Certified Planners (2021) AICP code of ethics and professional practice. https://www.planning.org/ethics/ethicscode/.

Amoore L (2016) Cloud geographies: Computing, data, sovereignty. Prog Hum Geogr 42(1):4–24. https://doi.org/10.1177/0309132516662147

Arnstein SR (1969) A ladder of citizen participation. J Am Inst Plan 35(4):216–224

Aguilera T, Artioli F, Colomb C (2019) Explaining the diversity of policy responses to platform-mediated short-term rentals in European cities: a comparison of Barcelona. Paris and Milan. Environ Plan 53(7):1689–1712. https://doi.org/10.1177/0308518X19862

Barandiaran X, Calleja-López A, Monterde A (2018) Decidim: political and technopolitical networks for participatory democracy. Decidim's project white paper.

Barns S (2014) Platform urbanism: The emerging politics of open data for urban management. American Association of Geographers Annual Conference, Tampa

Barns S (2015) Platform cities: The politics and practices of data-driven urbanism.

Barns S (2020) Platform urbanism: Negotiating platform ecosystems in connected cities. Palgrave Macmillan, Singapore

Benjamin, R. (2019) *Race After Technology: Abolitionist Tools for the New Jim Code.* Amsterdam University Press
Brossard M (2018) Artificial unintelligence. MIT Press, Cambridge
Buolamwini J, Gebru T (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Conference on fairness, accountability and transparency, January 2018. PMLR, p 77–91
Calzada I (2018) (Smart) citizens from data providers to decision-makers? The Case Study of Barcelona. Sustainability 10(9):3252. https://doi.org/10.3390/su10093252
Canadian Institute of Planners (2016) Code of professional conduct. https://www.cip-icu.ca/Careers-in-Planning/Codes-of-Professional-Conduct
Cardullo P, Kitchin R (2019) Being a 'citizen' in the smart city: Up and down the scaffold of smart citizen participation in Dublin, Ireland. GeoJournal 84(1). https://doi.org/10.1007/s10708-018-9845-8
CitizenLab (2023). "Your all in one platform for community engagement". https://www.citizenlab.co/platform-online-engagement-toolbox. Accessed 16 Jan 2023
Corbett E, Le Dantec CA (2019) "Removing barriers" and "creating distance": Exploring the logics of efficiency and trust in civic technology. Media Commun (lisboa) 7(3):104–113. https://doi.org/10.17645/mac.v7i3.2154
D'Ignazio C, Klein L (2020) Data feminism. MIT Press, Cambridge
Dickinson J, Diaz M, Le Dantec CA, Erete S (2019) "The cavalry ain't coming in to save us": Supporting capacities and relationships through civic tech. Proc ACM Hum-Comput Interact 3(CSCW). https://doi.org/10.1145/3359225
Einstein KL, Glick D, Godinez Puig L, Palmer M (2022) Still muted: The limited participatory democracy of Zoom public meetings. Urban Aff Rev. https://doi.org/10.1177/10780874211070494
Ellul J (1964) The technological society. Vintage Books
Eubanks V (2018) Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press, New York
Evans-Cowley JS (2010) (2010) Planning in the age of Facebook: the role of social networking in planning processes. GeoJournal 75:407–420. https://doi.org/10.1007/s10708-010-9388-0
Fields D, Bissell D, Macrorie R (2020) Platform methods: Studying platform urbanism outside the black box. Urban Geogr 41(3):462–468. https://doi.org/10.1080/02723638.2020.1730642
Fung A (2006) Varieties of participation in complex governance. Public Adm Rev 66(SI): 66–75.
Gagliardi D, Schina L, Sarcinella ML, Mangialardi G, Niglia F, Corallo A (2017) Information and communication technologies and public participation: Interactive maps and value added for citizens. Gov Inf Q 34(1):153–166. https://doi.org/10.1016/j.giq.2016.09.002
Gessa A, Sancha P (2020) Environmental open data in urban platforms: An approach to the big data life cycle. J Urban Technol 27(1):27–45. https://doi.org/10.1080/10630732.2019.1656934
Goetz EG, Williams RA, Damiano A (2020) Whiteness and urban planning. J Am Plan Assoc 86(2):142–156. https://doi.org/10.1080/01944363.2019.1693907
Goldsmith S, Kleiman N (2017) A new city O/S: The power of open, collaborative, and distributed governance. Brookings Institution Press, Washington. http://www.jstor.org/stable/https://doi.org/10.7864/j.ctt1vjqnwd
Gorwa R (2019) What is platform governance? Inf Commun Soc 22(6):854–871. https://doi.org/10.1080/1369118X.2019.1573914
Government of Canada (2020) Public engagement tools and resources. Privacy Council Services. https://www.canada.ca/en/privy-council/services/public-engagement-resources/content-analysis-consultations.html
Green B (2020) The smart enough city: Putting technology in its place to reclaim our urban future. MIT Press.
Gurran N, Phibbs P (2017) When Tourists Move In: How Should Urban Planners Respond to Airbnb? J Am Plan Assoc 83(1):80–92. https://doi.org/10.1080/01944363.2016.1249011

Höchtl J, Parycek P, Schöllhammer R (2016) Big data in the policy cycle: Policy decision making in the digital era. J Organ Comput Electron Commer 26(1–2):147–169. https://doi.org/10.1080/10919392.2015.1125187

Horgan D, Dimitrijević B (2019) Frameworks for citizens participation in planning: From conversational to smart tools. Sustain Cities Soc 48. https://doi.org/10.1016/j.scs.2019.101550

IAP2 (2018) IAP2 Spectrum of public participation. https://cdn.ymaws.com/www.iap2.org/resource/resmgr/pillars/Spectrum_8.5x11_Print.pdf

Janssen M, Charalabidis Y, Zuiderwijk A (2012) Benefits, adoption barriers and myths of open data and open government. Inf Syst Manag 29(4):258–268. https://doi.org/10.1080/10580530.2012.716740

Johnson PA, Scassa T (2023) Who owns the map? Data sovereignty and government spatial data collection, use, and dissemination. Trans GIS. https://doi.org/10.1111/tgis.13024

Johnson PA, Robinson P, Philpot S (2020). Type, tweet, tap, and pass: How smart city technology is creating a transactional citizen. Gov Inf Q 37(1). https://doi.org/10.1016/j.giq.2019.101414

Karner A, Brower Brown K, Marcantonio R, Alcorn LG (2019) The view from the top of Arnstein's ladder. J Am Plan Assoc 85(3):236–254. https://doi.org/10.1080/01944363.2019.1617767

Kemper J, Kolkman D (2018;2019) Transparent to whom? No algorithmic accountability without a critical audience. Inf Commun Soc 22(14):2081–2096. https://doi.org/10.1080/1369118X.2018.1477967

Kennedy, J (2015) How Digital Disruption Changed 8 Industries Forever. Silicon Republic, November 25, 2015.

Khan ZT, Johnson PA (2020) Citizen and government co-production of data: Analyzing the challenges to government adoption of VGI. Can Geogr. https://doi.org/10.1111/cag.12619

Kitchin R, Dodge M (2019) The (in)security of smart cities: Vulnerabilities, risks, mitigation, and prevention. J Urban Technol 26(2):47–65. https://doi.org/10.1080/10630732.2017.1408002

Kleinhans R, Van Ham M, Evans-Cowley J (2015) Using social media and mobile technologies to foster engagement and self-organization in participatory urban planning and neighbourhood governance. Plan Prac Res 30(3):237–247. https://doi.org/10.1080/02697459.2015.1051320

Knee J (2021) The Platform Delusion: Who Wins and Who Loses in the Age of Tech Titans. Portfolio/Penguin, New York

Lee A, Mackenzie A, Smith GJD, Box P (2020) Mapping platform urbanism: Charting the nuance of the platform pivot. Urban Plan 5(1):116. https://doi.org/10.17645/up.v5i1.2545

Legacy C (2017) Is there a crisis of participatory planning? Plan Theory 16(4):425–442. https://www.jstor.org/stable/26408371

Leszczynski A (2020) Glitchy vignettes of platform urbanism. Environ Plan D 38(2):189–208. https://doi.org/10.1177/0263775819878721

Mansell R (2015) Platforms of power. Intermedia 43(1):20–24. ISSN 0309–118X

Marvin S, Luque-Ayala A (2017) Urban operating systems: Diagramming the city. Int J Urban Reg Res 41:84–103. https://doi.org/10.1111/1468-2427.12479

Matheus R, Janssen M, Maheshwari D (2018) Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. Gov Inf Q 37(3). https://doi.org/10.1016/j.giq.2018.01.006

McIlwain CD (2020) Black software: The internet and racial justice, from the AfroNet to Black Lives Matter. Oxford University Press, New York

Nextdoor (2022) What is the Nextdoor public agencies program? https://help.nextdoor.com/s/article/what-is-the-nextdoor-for-public-agencies-program?language=en_US

O'Neil C (2016) Weapons of math destruction: how big data increases inequality and threatens democracy. Crown Publishers, New York

Paper presented at the Urban Studies Foundation Forum, University of Glasgow, 30 April 2015. https://www.slideshare.net/sarahbarns/platform-urbanism-the-politics-and-practices-of-datadriven-cities

Pasquale F (2015) The black box society: The secret algorithms that control money and information. https://doc.lagout.org/science/0_Computer Science/2_Algorithms/The Black Box Society_The Secret Algorithms that Control Money and Information %5BPasquale 2015-01-05%5D.pdf
Pina V, Torres L, Royo S (2017) Comparing online with offline citizen engagement for climate change: findings from Austria Germany and Spain. Gov Info Q 34(1):26–36. https://doi.org/10.1016/j.giq.2016.08.009
Robinson P (2022) Automation in municipal public consultation processes. In: Brandusescu, A, Reia, J (eds), Artificial intelligence in the city: Building civic engagement and public trust. Centre for Interdisciplinary Research on Montreal, McGill University, pp 19–20. https://doi.org/10.18130/9kar-xn17
Robinson P, Johnson PA (2021) Pandemic-driven technology adoption: public decision makers need to tread cautiously. Int J E-Plan Res 10(2):59–65. https://doi.org/10.4018/IJEPR.20210401.oa5
Robinson P, Biggar J (2021) Seeing the city as a platform: Is Canada's smart cities challenge a step in that direction? In: Zwick A, Spicer, Z (eds) From the platform economy to the smart city: Technology and the transformation of urban policy and governance. McGill-Queen's University Press
Robinson P, DeRuyter M (2016) I "like" you, you make my heart twitter, but…" Reflections for urban planners from an early assessment of social media deployment by Canadian local governments. In: Thomas, R (ed) Planning Canada: a case study approach. Oxford University Press, Don Mills, pp 260–268
Rosenblat A (2018) Uberland: how algorithms are rewriting the rules of work. University of California Press
Schweitzer LA, Afzalan N (2017) 09 F9 11 02 9D 74 E3 5B D8 41 56 C5 63 56 88 C0: four reasons Why AICP needs an open data ethic. J Am Plann Assoc 83(2):161–167
Sieber RE, Robinson P, Johnson PA, Corbett JM (2016) Doing public participation on the geospatial web. Ann Am Assoc Geogr 106(5):1030–1046. https://doi.org/10.1080/24694452.2016.1191325
Silva P, Tavares AF, Silva T, Lameiras M (2019) The good, the bad and the ugly: three faces of social media usage by local governments. Gov Inf Q. https://doi.org/10.1016/j.giq.2019.05.006
Streich B (2018) Subversive urban planning: Alternative forms of urban action in the knowledge society. ISBN 978-1986756617
Wiig A (2016) The empty rhetoric of the smart city: from digital inclusion to economic promotion in philadelphia. Urban Geogr 37(4):535–553. https://doi.org/10.1080/02723638.2015.1065686
van der Graaf S, Ballon P (2019) Navigating platform urbanism. Technol Forecast Soc Change 142:364–372. https://doi.org/10.1016/j.techfore.2018.07.027
Van Doorn N (2019) A new institution on the block: On platform urbanism and Airbnb citizenship. New Media Soc 22(10). https://doi.org/10.1177/1461444819884377

Part II
Mobility Futures

Chapter 6
Shared Micro-mobility: A Panacea or a Patch for Our Urban Transport Problems?

Zhenpeng Zou

Abstract Shared micro-mobility, including station-based bike-sharing and dockless bike-/scooter-sharing, experienced phenomenal growth in the past decade in cities across the globe. It is low traffic impact, eco-friendly, and associated with a healthy lifestyle. Cities see it as a viable solution to solve issues related to congested, polluted, and auto-centric urban transport. In this chapter, I overview the history of shared micro-mobility. I then broadly summarize the existing research on shared micro-mobility systems around the world and explore how shared micro-mobility has transformed urban transport for its users. Using an empirical case from the city of Brisbane, Australia, I demonstrate the usage and limitations associated with shared micro-mobility trip big data. Lastly, I narrate two possible scenarios of the shared micro-mobility future. I conclude the chapter with a call for collaborations between cities, vendors, and researchers to make shared micro-mobility work for our future urban transport.

Keywords Shared mobility · Micro-mobility · e-scooters · Clustering analysis · Data analytics

6.1 Shared Micro-mobility: The New Kid on the Block

Open the app, search for an available vehicle nearby, walk to it, scan the QR code on the app, unlock the vehicle, start a trip, travel, park at the destination, and end a trip. I just walked you through a standard shared micro-mobility trip—no muss, no fuss. Micro-mobility, supported by light-weighted, low-speed (<25 km/h or 15 mph), human-/electric-powered fleets (e.g., bike, scooter, self-balancing board, and segway), has become an emerging and popular mode of mobility for short-distance travel (<5 km) in the urban environment (Oeschger et al. 2020; Orozco-Fontalvo et al. 2022). Shared micro-mobility, as the name suggests, is a shared mobility system for micro-mobility. A contemporary shared mobility system can be characterized by its

Z. Zou (✉)
School of Earth and Environmental Sciences, The University of Queensland, St Lucia, Australia
e-mail: zhenpeng.zou@uq.edu.au

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_6

reliance on online platforms (i.e., mobility vendors), mobility as a service (MaaS), and an on-demand marketplace that matches demand with supply in real time. Ride-hailing services provided by transportation network companies (TNCs) like Uber or DiDi, car-sharing services provided by car rental companies, and shared micro-mobility services provided by cities and private vendors are all broadly umbrellaed under the shared mobility concept. The global micro-mobility market surpassed $100 billion in 2021 with a 10 percent non-ownership market share (Lang et al., 2022). In particular, the shared micro-mobility market is projected to grow 10–30% over this decade (Lang et al., 2022). As the new kid on the block of urban transport, shared micro-mobility is reviving a niche mobility option in a heavily auto-oriented urban world.

6.1.1 The Three Generations of Shared Micro-mobility

Shared micro-mobility systems around the world went through a three-stage evolution based on the significant advancements in technology and popularity/influence:

The 1960s—early 2000s (ad hoc, city-operated bike-sharing) While the first verifiable bicycle—*draisine*—was introduced to the world back in the 1810s in Germany (Herlihy, 2004), the first public bike-sharing system—*Witte Fietsenplan* ("white bike plan")—only emerged 150 years later in Amsterdam, The Netherlands, out of an idealistic mission to take the street back from "the gaudiness and filth of the authoritarian car" (van der Zee, 2016). Ahead of its time, the system was merely a grassroots response to the urban transport crisis. However, the idea eventually incubated a global micro-mobility initiative decades later. In the 1970s, public bike-sharing systems were established in several European cities, such as *vélo* in La Rochelle, France, was developed as an institutional solution to address urban sprawl and auto-dependency (Hure & Passalacqua, 2017). In 1995, the Danes launched the world's first large-scale bike-sharing system, *Bycykler København*, with 1,100 bicycles locked and placed all over Copenhagen. Of course, this system predated smartphone apps such that a coin-deposit system was utilized for payment.

In addition, a docking-station model was adopted for fleet rental and return (Shaheen et al., 2010). This became a prototype for the contemporary station-based bike-sharing (SBBS) systems. The drawback of a SBBS system is obvious: fleets are subject to vandalism and theft. Moreover, coin-deposit payment is not user-friendly. In the early 2000s, smart technologies (e.g., mag-stripe cards) for check-out and check-in of a bike, payment, and theft deterrents were integrated into bike-sharing systems (Abduljabbar et al. 2021)—a critical upgrade towards the modern bike-sharing system. Nevertheless, it is reasonable to state that before the 2010s, bike-sharing was a niche in the urban transport market.

The late 2000s—early 2010s (app-based, station-based, city-led bike-sharing programs). As smartphones have become ubiquitous to the average person and an ecosystem of smartphone apps has been built to facilitate every aspect of our daily life since the late 2000s, it was only a matter of time before shared micro-mobility evolved into the 2nd generation. In this stage, many existing public bike-sharing programs were upgraded to be readily run via a smartphone app. The major advantage of an app-based system is the integration of searching, booking, navigation, and payment, all under one platform instead of having isolated platforms each perform a single function.

Consequently, bike-sharing systems quickly diffused into major cities around the world. In Europe, a significant number of SBBS schemes were launched between 2007 and 2012 (Parkes et al., 2013). In North America, the bike-sharing movement finally caught up: BIXI Montréal was the first large-scale public bike-sharing system launched in 2009. In the U.S., public bike-sharing expanded from four programs in 2010 to 55 programs in 2016 with annual trips growing 9-folded from 0.3 million to 2.8 million in the same time frame (NACTO, 2017). In Australia, Brisbane's CityCycle and Melbourne's Bike Share debuted in 2010. Asia by and large missed the opportunity to embrace shared micro-mobility in this stage, despite several countries that have a strong bike culture traditionally, such as China, Japan, and Vietnam.

Another key feature of bike-sharing systems in this generation is the use of a docking station for rental and return. The semi-flexible docking-station model has advantages and drawbacks: It allows for more orderly fleet dispatching and rebalancing, but the geographical coverage of stations is limited due to the non-trivial construction cost, land use impact, and capacity limit of a docking station (Chen et al., 2020). Lastly, the bike-sharing programs in this stage were mostly launched and operated by cities with some partnerships with private vendors (Shaheen et al., 2010). While the city-led model worked well in the early-day diffusion of shared micro-mobility, its limited ability to grow with the market was magnified when cities hesitated to invest a significant amount of budget to expand their bike-sharing program.

The late 2010s—present (platform-mediated, dock-less, shared micro-mobility programs). In 2016, a number of Chinese dock-less bike-sharing (DBS) start-ups (e.g., Mobike and Ofo) rolled out their product both domestically and internationally and achieved a visible expansion backed by venture capital investment (Zou et al., 2020). It marked the beginning of a new generation of shared micro-mobility. The dock-less eco-system further cannibalized the previous generation's market share when a new player—e-scooters—was added to the shared micro-mobility family. In the U.S., e-scooter sharing became the predominant mode of shared micro-mobility, as compared to SBBS and DBS, in 2019 (NACTO, 2020). In Canada, Australia, New Zealand, Europe, and Asia (e.g., Singapore), dock-less e-scooter sharing emerged around the same time and proliferated at an impressive speed as well.

The dock-less operation model relaxes both the geographical confinement and the fleet capacity limit of a docking station. More importantly, a private-public partnership business model between cities and vendors significantly motivates capital to invest in shared micro-mobility (Zou, 2021). Shared micro-mobility vendors bring in operational and economic efficiency that helps accelerate the mass deployment of thousands of e-bikes and e-scooters on city streets. Although shared micro-mobility took its toll during the Covid-19 pandemic, industry statistics from North America suggest that the market has quickly recovered—even surpassing its pre-pandemic level (North America Bikeshare & Scootershare Association (NABSA), 2022). At this stage, cities and the transport industry are taking a serious look at this new kid on the block: Is shared micro-mobility going to take over a substantial share of the urban transport market? What does a shared micro-mobility vision look like for cities?

6.1.2 Contribution and Disruption to Cities

The numbers do not lie: Shared micro-mobility's popularity is quickly rising. Cities are gearing up to deploy more bikes and scooters in the street, build more docking stations, and permit more service vendors. They must see the desirability of shared micro-mobility to potentially remedy the wicked problems in urban transport, such as congestion and the steady decline in public transport ridership. In Brisbane—the city where I live and work, micro-mobility is heavily promoted by the Brisbane City Council. The Council highlights that micro-mobility brings an array of benefits, including an affordable mobility option, convenience to tourists and residents, promoting sustainability in urban transport, spurring economic growth through job creation and local business visitations, and creating a safer, pedestrian-/cyclist-friendly built environment (the Brisbane City Council, 2020). Because the fleets are associated with a healthy, green, low-traffic-impact image a generally favorable sentiment towards shared micro-mobility is found in qualitative evidence across global cities (Mitra & Hess, 2021; Bakker, 2018).

On the other hand, the shared micro-mobility vision is overshadowed by a few outstanding issues. In its early days, the issue of dock-less fleets cluttering streets/sidewalks was pervasive (Burtina et al., 2020). If the war of who owns sidewalks and curbsides reflects spatial mismanagement, then a system-wide mismanagement of fleet supply could cause an even more undesirable consequence: For instance, China's bike-sharing mania in 2017–2018 resulted in thousands of damaged and abandoned bikes piling up and rotten in the so-called 'bike graveyard'. Production and supply of fleets, pumped by ill-considered investors rushing into this emerging market, dramatically outpaced the actual demand for shared micro-mobility in multiple Chinese cities. Eventually, it led to huge piles of fleet debris (Taylor, 2018). Additionally, shared micro-mobility was, and still is, scrutinized for safety issues. Particularly, e-scooters can travel at a comparable speed to cars in low-speed (~20 mph) city streets. When e-scooter riders share the roadway with drivers, traverse congested intersections, or ride on poorly lit roads at night, they run into the risk of a crash incidence that may result in an injury to a varied degree of severity. On the other hand, shared micro-mobility riders may also become a safety hazard/ nuisance to pedestrians and residents (Aman and Smith-Colin 2021). Moreover, researchers question whether shared e-scooters generate any environmental benefit based on simulation outcomes, suggesting that their environmental impact is highly sensitive to the emission from charging and the actual private car travel distance being replaced over the life cycle of an e-scooter (Hollingsworth et al. 2019). In a case where e-scooters have a short life span (less than one year) and an insufficient number of auto trips are being replaced the emission level could be higher than a baseline without any shared e-scooter.

This chapter focuses on two important questions: What impacts shared micro-mobility has had and may have on our urban transport? As micro-mobility vendors continue upgrading the hardware (e.g., battery life, durability, maneuverability, safety, and anti-theft features of an e-bike/e-scooter) and the software (e.g., GPS accuracy, geo-fencing accuracy, and facial recognition of a user), we can foresee that cities will either continue expanding their shared micro-mobility program or jump on the bandwagon if they do not have one currently. However, how big a promise can shared micro-mobility fulfill in terms of building a more efficient, equitable, and eco-friendly urban transport future?

To answer the questions, the rest of this chapter is centered around three activities: In Sect. 6.2, I summarize the existing research evidence on how shared micro-mobility has transformed/ is transforming our cities; In Sect. 6.3, I use big shared micro-mobility trip data to empirically demonstrate what the data does and doesn't tell about the usage of this novel mobility in our urban environment; In Sect. 6.4, I create two qualitative narratives on dichotomous micro-mobility futures; In Sect. 6.5, I conclude the chapter by discussing what cities, vendors, and researchers should jointly work on to achieve a sustainable micro-mobility future that we have been envisioning since the idea was born half a century ago.

6.2 Shared Micro-mobility Transforming Cities: The Research Landscape

With public shared micro-mobility programs being implemented in cities around the world since the 2010s, research that characterizes shared micro-mobility business, operation, travel pattern, and user profiles have also taken off across different disciplines. In this section, I pay attention to studies that try to empirically understand, uncover, and underpin the extent to which shared micro-mobility has reshaped our urban economy, sustainability, accessibility, and lifestyle. By no means is this review exhaustive or systematic as (1) innovations in technology, business, and operation of shared micro-mobility continue to improve this new product, and (2) localized knowledge dominates general knowledge in this emerging research area as each local market offers unique practices and perspectives.

6.2.1 Shared Micro-mobility Transforming the Urban Economy

Transport is the backbone and vessel of the urban economy. Shared micro-mobility transforms the urban economy in the digital, location-based direction. It benefits the urban economy by improving transport efficiency. As many cities are faced with severe peak-hour congestion, shared micro-mobility (e-bikes and e-scooters) offers a faster urban mobility solution compared to driving for short-distance commutes (McKenzie, 2020). In addition, empirical studies (Ma et al. (2015) for SBBS and Yang et al. (2019) for DBS) suggest that bike-sharing can be integrated with public transport in a first-/last-mile setting and boost transit ridership. It is yet to be empirically concretized that e-scooter sharing complements public transit, although conceptually it is plausible (Zuniga-Garcia et al., 2022).

The economic efficiency argument is not without limitations. In terms of replacing car travel, survey evidence points to divergent possibilities of either a substitutional (especially for short-distance trips of <2 miles) or an insignificant relationship (due to auto-dependency) between cars and e-scooter sharing (Wang et al. 2022). Public bike-sharing trips may also replace short or unlinked transit trips, which may adversely affect transit ridership (Campbell and Brakewood 2017), especially during the Covid-19 pandemic (Teixeira and Lopes 2020).

Besides improving transport efficiency, shared micro-mobility and its associated infrastructure (e.g., docking stations) are seen as part of the millennial economy, just like craft-beer shops and corner cafés, that reshapes the local economy and uplifts previously unattractive neighborhoods (Hyra 2016). Indeed, cluttered sidewalks may be seen as an eyesore, but a well-balanced docking station/parking corral with bright-colored, fresh-looking e-bikes and e-scooters is place-making in itself (Buehler & Hamre, 2016). The flip side of the aesthetic appeal of shared micro-mobility is the potential gentrification of an affordable community (Leszczynski and Kong 2022).

6.2.2 Shared Micro-mobility Transforming Urban Sustainability

It was the congested, auto-oriented urban transport system that impregnated the original idea of shared micro-mobility. Naturally, shared micro-mobility is seen as a solution to reduce auto-dependency and mobile source emission. Empirical studies using American and Chinese public bike-sharing cases have demonstrated its environmental benefits. Hamilton & Wichman (2018) identified a causal, positive impact of public bike-sharing in Washington D.C. on congestion reduction (by as much as 4%). He et al. (2020) demonstrated that bike-sharing usage in four U.S. cities is highly sensitive to fuel costs for driving, indicating that micro-mobility can substitute short-distance & short-duration driving trips. Consequently, a small but significant reduction in vehicle distance traveled is translated into a small reduction in CO_2 emission. In another case from Shanghai, China, Zhang and Mi (2018) quantified emission reduction as a result of the proliferation of bike-sharing.

On the other hand, whether shared micro-mobility can realize its promises of sustainability largely depends on (1) the lifecycle emission, especially for e-bikes and e-scooters, and (2) the level of substitution between shared micro-mobility and auto travel. Luo et al. (2019) compared the life-cycle emission of DBS and SBBS systems: The former's emission level is 82% higher than the latter in a life cycle. Hollingworth et al. (2019) suggest that a short lifespan or a low battery efficiency of a scooter would significantly limit its environmental benefit. Furthermore, mixed results on the substitution between shared micro-mobility and car travel were identified across different countries/systems (Fishman et al. 2014). To accurately assess the environmental benefit of shared micro-mobility, we need to consider both the emission reduction through the vehicle travel distance being substituted and the emission accumulated throughout a fleet's life cycle (fleet production, the battery recharges, rebalancing, and recycling).

6.2.3 Shared Micro-mobility Transforming Urban Accessibility

Shared micro-mobility is seen as a convenient mobility option to improve people's accessibility to various opportunities and amenities. Public transport operates on fixed routes with a fixed timetable, which means there are blind spots when and where it is unavailable to serve those who lack mobility. Public bike-sharing can readily cover such blind spots, providing the transit-dependent population with a first-/last-mile solution or directly substituting short transit trips (Kong et al., 2020). In addition, many cities designate equity priority areas, where many low-income and low-mobility households are guaranteed to have a fair share of micro-mobility fleets and infrastructure (Zou, 2021). Cities also implement financial assistance programs to make shared micro-mobility either affordable or free for the most economically disadvantaged individuals (Riggs et al., 2021).

One limitation of spatial access is that it does not directly translate into usage. Membership subscription amongst socially disadvantaged residents tends to be lower than others, although they are inclined to use the service more often upon joining membership (Qian & Jaller, 2020). Furthermore, even if the spatial coverage of fleets may be adequate in low-income neighborhoods, significantly fewer trips tend to be generated in such areas, suggesting a disparity in the actual usage (Frias-Martinez et al. 2021).

6.2.4 Shared Micro-mobility Transforming Urban Lifestyles

Finally, shared micro-mobility promotes an active, healthy lifestyle that is aligned with the public health mission for cities. The health benefit associated with bike-sharing, such as lowering the obesity rate, substantially outweighs the risk of a higher morbidity/mortality rate associated with traffic accidents (Woodcock et al. 2014). Empirical evidence suggests that bike-sharing is associated with more physical activity and lower rates of obesity (Xu, 2019; Stahley et al. 2022). However, there lacks sufficient evidence to prove the health benefit of e-scooter sharing due to its novelty and the fact that riding an e-scooter requires little physical effort.

The image associated with shared micro-mobility—high-tech, green, affordable, and active—softens disputes and controversies caused by its early-day oversupply, mismanagement, and safety concerns. Policymakers and governments see the potential of shared micro-mobility to transform the urban environment. Nonetheless, there exist significant knowledge and information gaps between a pilot shared micro-mobility program and mass adoption—where the high-demand areas and corridors are, how to balance demand and supply, who uses or refuses shared micro-mobility, what the traffic impacts are, etc. Such knowledge must be gained through evidence-based research specific to a local shared micro-mobility program. In the next section, I will showcase an empirical work that explains what typical shared micro-mobility data can and cannot inform decision-makers about the travel and operation patterns of shared micro-mobility.

6.3 What Does and Doesn't The Data Tell? An Empirical Demonstration

In the age of mobility big data, trip data is arguably the most available data source for shared micro-mobility. Through web-scraping techniques and partnerships with private vendors, researchers can either self-collect trip data or receive data from service providers. Combining spatial and temporary trip information with secondary data (e.g., weather, points of interest, population statistics), researchers can make reasonable inferences about the shared micro-mobility users' travel behavior. In my case, I will interpret trip patterns in Brisbane, Australia, using data-driven empirical techniques.

6.3.1 Study Area

Brisbane serves dual functionalities: a business city and a tourism city. As the capital city of the State of Queensland, Brisbane has a 1.3 million population in the labor force (Australian Bureau of Statistics (ABS), 2022), which generates sizeable commuting trips. The city also possesses enormous tourism resources: Brisbane is known for its proximity to world-class beaches on the Gold Coast and the Sunshine Coast, as well as its tourist attractions in the central business district (CBD) and along the Brisbane River. In terms of urban form, Brisbane is a classic monocentric city with a predominant employment/activity center in the CBD. Population density and building density quickly decay outside of a 3km buffer from the CBD, resulting in auto-oriented suburbs.

Demographic-wise, Brisbane has a diverse immigrant population (31.7% oversea born population in 2021(ABS, 2022) with a strong presence of the Asian community. Climate-wise, Brisbane has a year-round subtropical warm and humid weather, which favors cycling activities.

6.3.2 Data

I processed shared micro-mobility trip data received from a Singaporean-based vendor—Neuron, who entered the Brisbane market in 2019. For demonstration purposes, I analyzed trips between August 30 and September 26, 2020—the State of Queensland had its borders (interstate and international) closed against the spread of the Covid-19 Delta variant at the time. Essentially, the micro-mobility trips were taken by Queenslanders, if not Brisbane locals, during the demonstration period. A total number of 64,353 trips were included in the analysis.

I spatially joined trip origins and destinations with the following spatial features: points of interest (POIs), including educational institutes, financial services, healthcare facilities, hotels, parks, public buildings, stores and shopping centers, sports venues, sightseeing spots, restaurants and bars, and churches, from OpenStreetMap; transport infrastructure locations, including bus hubs, train stations, and ferry terminals, from general transit feed specification (GTFS) data provided by the regional transit authority; and roadway classification (major roadways, minor roadways, and bikeways) from OpenStreetMap. A 50m buffer for various (POIs) and a 10m buffer for the street network were drawn. Then, the number of spatial features within which a trip's origin and destination fell inside was counted.

In addition, I merged 0.5-hourly archived Brisbane weather data (temperature, humidity, wind speed, and precipitation) from Weather Underground (https://www.wunderground.com/) with trip data by a trip's starting timestamp. Lastly, trip characteristics themselves were considered in the analysis, such as time of day, day of week, distance, duration, and loop trip—where its origin-destination distance is shorter than 200m and 1/3 of its network distance.

6.3.3 Building a Typology of Shared Micro-mobility Trips

With rich information from multiple data sources, I was able to explore what big trip data can tell us about travel behavior. I applied the same analytical framework from my previous work (Zou, 2021), where I built a typology of e-scooter trips using a K-means clustering technique in Washington, D.C. based on the spatiotemporal features of an e-scooter trip. Methodologically, building a 'typology' of shared e-scooter trips draws inspiration from the transit user segmentation studies (e.g., Grise and El-Geneidy, 2018) and bike-sharing spatiotemporal trip characterization studies (e.g., Zhou, 2015). Based on the qualitative description of natural and built environment factors that distinguish trips within one cluster from other clusters, I can make inferences about unique/common travel behavior associated with shared e-scooters in Brisbane.

A flow chart conceptualizing the empirical framework is provided as follows in Fig. 6.1.

Notice that ideally user data should be accompanied for a better understanding of individual users' travel behavior. However, such information is not available in the dataset (Hence the dashed box of 'User data' in Fig. 6.1). Empirically, the following bullet points summarize key processes of building the trip typology:

- The following features were run into the clustering analysis: 16 spatial features related to trip origin, 16 spatial features related to the trip destination, 10 features related to a trip's temporal information, 2 features related to trip distance and duration, as well as 4 features related to the weather conditions at the time when a trip was taken
- An elbow test was run to determine the optimal number of clusters, which is 15
- 9 out of the 15 clusters have a decent size (>2,000 out of 64,353 trips): a qualitative summary of the distinct characteristics of each cluster was created

For demonstration, the qualitative summary for Cluster 11—'weekend riverside recreational trips' (n = 5,316) is presented in Fig. 6.2: Out of the 48 spatiotemporal features, only the features with a significantly higher cluster mean statistic (the blue series) than the sample mean statistic (the gray series) are visualized. The number of point of interest (POI) buffered areas (typically 50 meters) an average shared e-scooter trip's origin/destination falls inside, an average trip's weekday/weekend status, whether a trip is a loop trip or long duration trip are significant factors that differentiate trips within Cluster 11 and the entire sample. As illustrated in Fig. 6.2, on average, trips in Cluster 11 were more likely to start and end near public buildings (e.g., library and city hall), sightseeing locations, etc. In addition, these trips were also more likely to be long, loop trips taken on a weekend day along bikeways, hinting at recreational rather than utilitarian usage.

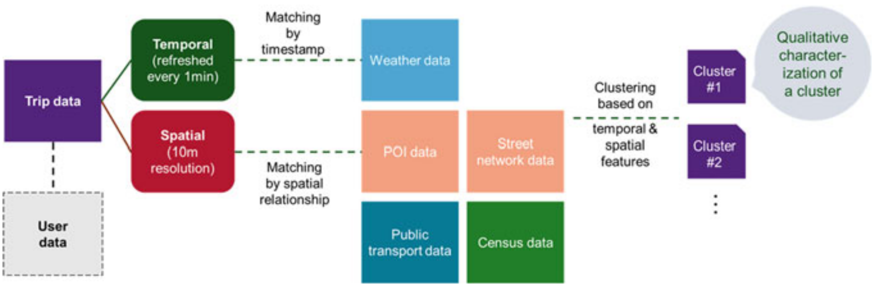


Fig. 6.1 The flow chart of the empirical framework

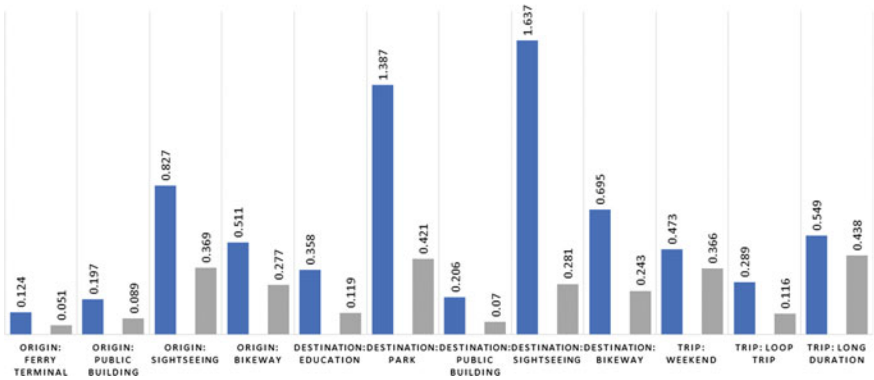


Fig. 6.2 Summary of the distinct characteristics for Cluster 11 (the blue series: cluster mean statistics; the gray series: sample mean statistics)

Figure 6.3 maps trip origins (top) and destinations (bottom) for Cluster 11. The spatial distributions of origins and destinations confirm the distinct spatial characteristics summa: A significant number of trips started and ended along the Brisbane River in the CBD and South Bank with proximity to various POIs and on riverfront bike trails.

Clustering is a powerful data analytic tool to recognize shared micro-mobility trips of similar spatiotemporal dynamics. Taking advantage of real-time big data with a range of features, we can distinguish and explain heterogeneous mobility patterns. The geospatial and temporal clusters of micro-mobility trips can then inform stakeholders (vendors, transport authorities, city planners, and policymakers) about the points of interest that generate high trip demand, the corridors that could benefit from new/improved bike infrastructure, and the specific hours in a day in a week to manage micro-mobility traffic.

6.3.4 People: The Missing Puzzle

Nonetheless, a major factor is missing from the analysis: the people—specifically, the riders who rode e-bikes and e-scooters. Are they visitors or residents? How often do they bike/scoot? Is a certain trip type (as indicated by 'clusters') attached to certain user demographics?

In the research space, it is not an uncommon practice to anonymize user information in passive trip data due to privacy concerns (Cottrill 2020). Particularly in the age of big data, researchers hope to capitalize on the geographic and temporal granularity of big trip data and explore travel behavior at a fine-grained scale. Had we possessed a range of user information (e.g., membership status, user ID, and basic sociodemographic characteristics), we should be able to fully characterize (1) shared micro-mobility user behavior, (2) the existing market penetration amongst different

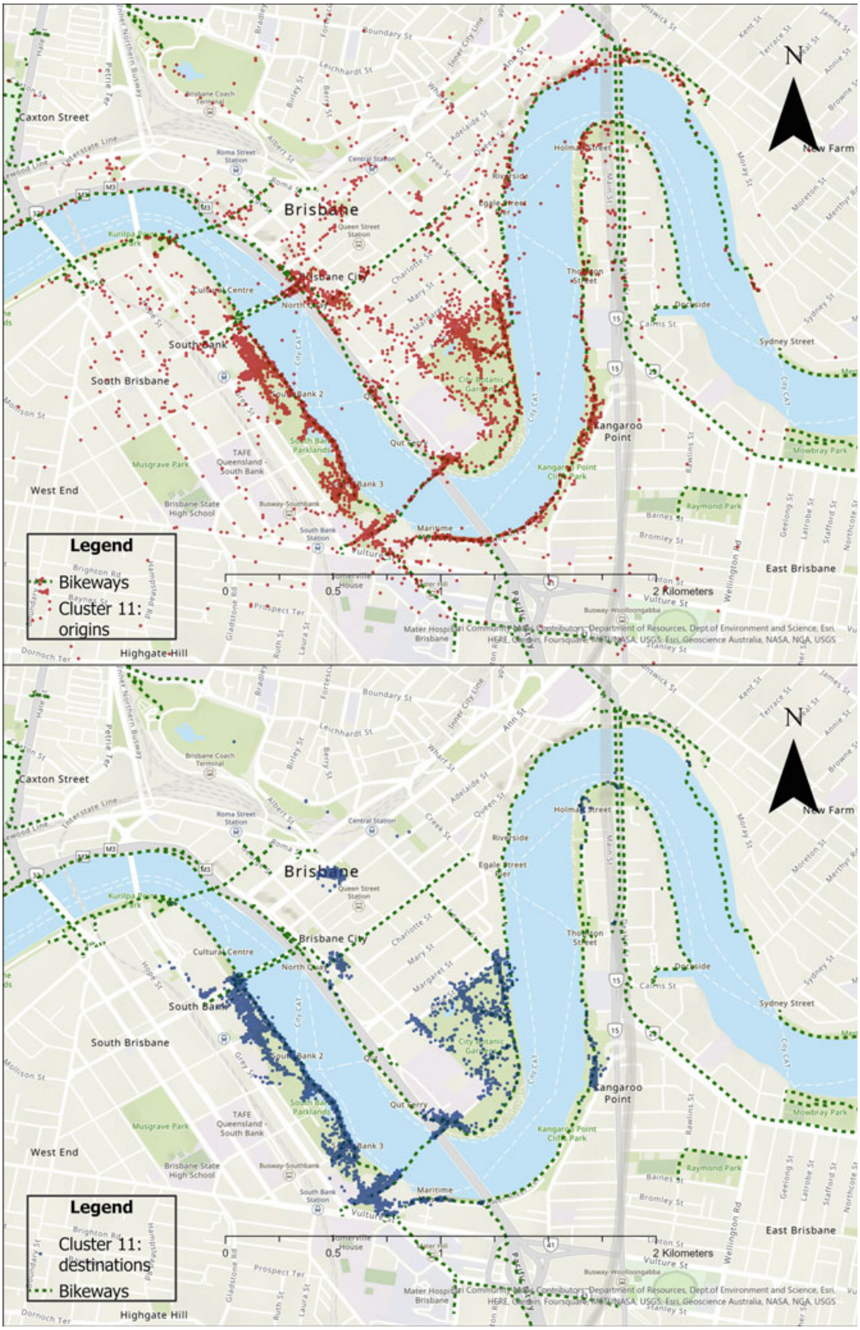


Fig. 6.3 The spatial distributions for trip origins (top) and destinations (bottom) for Cluster 11: Weekend riverside recreational trips

sub-population groups, and (3) differentiation between casual and membership users. On the other hand, we still need to collect data on people to understand why they use or refuse to use shared micro-mobility occasionally and/or frequently.

The empirical work makes the case that data is a powerful friend which informs high-level decisions to manage or expand a shared micro-mobility program, such as the system-wide fleet quota, effective and equitable spatial deployment of fleets, critical corridors to add new bikeways, intersections to accommodate multimodality, and major activity areas in need for parking management (e.g., geofencing/parking corrals). Significant data gaps when it comes to 'people' should also be acknowledged. A trustworthy partnership among the city, private and public vendors, and researchers need to be established to start incorporating the user (and non-user) dimension into analyses on shared micro-mobility travel behavior whilst mindfully protecting user data privacy. Anonymity and aggregation are viable solutions, but consultation is necessary to achieve a balance between meaningful analyses and the non-disclosure of personal information.

6.4 Two Scenarios of Shared Micro-mobility: A Panacea or a Patch for Urban Transport Problems?

Cities champion shared micro-mobility. Vendors expand the shared micro-mobility market. Researchers facilitate these two decision-makers with data-driven, evidence-based studies. Do the inputs from all three converge and point to a shared micro-mobility vision? If so, what does it possibly look like, and what it means to our problematic urban transport? While I do not have an answer, I could imagine two scenarios of the future shared micro-mobility:

6.4.1 Scenario 1: A Shared Micro-mobility Paradise

This scenario sketches a fresh blueprint of an urban transport ecosystem centered around micro-mobility as opposed to a minor increment from the status quo. For this scenario to happen, several factors need to work together:

Firstly, we need a thriving shared micro-mobility market. On the *supply* side, many vendors, not just a few start-ups, will enter the shared micro-mobility market. They provide adequate services on varieties of products and a tiered pricing scheme that suits user groups of varied mobility needs and income levels. On the *demand* side, shared micro-mobility can meet most of our travel demands, including daily commuting, recreation, shopping, social and nightlife, as well as out-of-town visits. For longer-distance trips, micro-mobility can provide first-/last-mile connectivity to public transit.

The *built environment* must strongly favor micro-mobility in this paradise. On the *infrastructure* side, Bike-friendly infrastructure is installed in major activity centers, corridors, and transit hubs, such as bike superhighways, off-street bike trails, on-street protected bike lanes, bike/scooter corrals, and fast-charge stations. A high-density, mixed-use, job-housing-balanced, polycentric *urban form* is desirable in this scenario. Commuting trips and non-commuting trips can be accomplished within 30 minutes and 15 minutes, respectively.

Public policy also plays a critical role in promoting micro-mobility. Cities permit vendors to operate on a generous fleet quota. Car-free zones are designated for pedestrians and micro-mobility riders in activity centers. Auto traffic speed is capped at 25mph on local and minor arteries to protect micro-mobility riders' safety. Periodical data sharing between cities, vendors, and researchers is arranged using a standard, anonymous format.

Suppose all the prerequisites are met, then we shall see a significant mode shift from private vehicles to (1) shared micro-mobility for short- & middle-distance trips, and (2) a 'shared micro-mobility + public transport' hybrid mode for long-distance trips. In the long run, this shared micro-mobility heaven will have significantly less traffic congestion, less pollution, better traffic safety, a more active lifestyle, and improved health outcomes for all citizens.

6.4.2 Scenario 2: When the Hype is Over

If Scenario 1 portrays an ideal world for shared micro-mobility, then Scenario 2 depicts a much drearier shared micro-mobility future, where cities no longer expand their program and the other factors are not much different from the status quo:

The shared micro-mobility market remains a niche in this scenario. On the *supply* side, only a limited number of vendors and fleets are permitted to operate in the core urban areas. On the *demand* side, residents and visitors of certain demographic and socioeconomic characteristics would use shared micro-mobility services, such as young, physically active, high-income, and car-less tourists and residents who live in the urban core.

The built environment favors private cars, instead of pedestrians and cyclists. On the *infrastructure* side, a marginal expansion/improvement of bike infrastructure is implemented. In terms of *urban form*, cities remain monocentric, sprawling in the periphery, and auto-centric.

Lukewarm promotion of shared micro-mobility is enacted by cities. In addition, cities offer minimal incentives for vendors to enter the market or for residents and visitors to try out micro-mobility services. Limited collaboration between cities and researchers is achieved to understand travel demand and travel behavior associated with new mobility.

We would expect a negligible mode shift from private vehicles to shared micro-mobility in this scenario. What is worse, with ongoing population growth and suburbanization auto-dependency may further aggravate. In this scenario, only a fringe user group benefits from shared micro-mobility. However, the well-being of the general public either remains unchanged or worsens from the status quo.

Of course, transport technologies advance beyond our prediction and imagination. Emerging mobility options, such as electric and autonomous vehicles, and alternative travel methods, such as teleworking and online delivery, may fundamentally alter our future urban transport system. Shared micro-mobility may not be a solution but a transition. While our society should keep up with technological advancement, we should keep in mind that efficiency is not the sole metric that matters. Green, active, and affordable micro-mobility offers multitudes of broader social benefits beyond economic efficiency.

6.5 Conclusion: Making Shared Micro-mobility Work for Cities

In this chapter, I walked through the origin and evolution of shared micro-mobility. I broadly introduced the current research findings on how shared micro-mobility is transforming cities. I also empirically demonstrated the power of big data associated with shared micro-mobility, along with its limitations where the human factor is missing. Finally, I described two possible shared micro-mobility futures moving forward. Although researchers deliver mixed messages, vendors sometimes mismanage, and cities see it as both a friend and a foe, we can reach a consensus that shared micro-mobility does more good than harm to cities.

Furthermore, a shared micro-mobility paradise is not a delusion – do you know it has the nickname 'Copenhagen'? If car-oriented cities around the world scrap their lukewarm bike and pedestrians master plans and follow the playbook of building a micro-mobility heaven (inspired by the Danes!), then our cities will be much greener, safer, more pleasant places to live in.

As a researcher in shared micro-mobility myself, I believe we need to make our research outcomes visible to cities and vendors. In return, they will seek collaboration with us by providing much-needed data and funding capacities. Together, we can ally to re-draw the blueprint of urban transport—with a little help from shared micro-mobility!

Acknowledgement The author would like to acknowledge Neuron's contribution to providing Brisbane e-scooter sharing trip data for research purposes. The empirical analysis solely reflects the author's own interpretation and is not subject to Neuron's company values or opinions. This research is not funded externally by any grant. All images in this chapter are produced by the author.

References

Abduljabbar, R. L., Liyanage, S., & Dia, H. (2021). The role of micro-mobility in shaping sustainable cities: A systematic literature review. *Transportation Research. Part D, Transport and Environment*, 92, 102734. https://doi.org/10.1016/j.trd.2021.102734

Aman, J. & Smith-Colin, J. (2021). Leveraging Social Media to Understand Public Perceptions toward Micromobility Policies: The Dallas Scooter Ban Case. *Findings*.

Australian Bureau of Statistics (ABS). (2022). Greater Brisbane 2021 Census All Persons QuickStats. Available at: https://www.abs.gov.au/census/find-census-data/quickstats/2021/3GBRI

Bakker, S. (2018). Electric Two-Wheelers, Sustainable Mobility and the City. Chapter 6 in *Sustainable Cities: Authenticity, Ambition and Dream*. http://dx.doi.org/https://doi.org/10.5772/intechopen.81460

Bliss, L. (2017). Cruising a Superhighway Built for Bikes. In Bloomberg, CityLab Transportation. https://www.bloomberg.com/news/articles/2017-06-22/this-dutch-cycling-superhighway-connects-commuters

Buehler R, Hamre A (2015) Business and Bikeshare User Perceptions of the Economic Benefits of Capital Bikeshare. Transportation Research Record 2520(1):100–111. https://doi.org/10.3141/2520-12

Butrina, P., Le Vine, S., Henao, A., Sperling, J., & Young, S. E. (2020). Municipal adaptation to changing curbside demands: Exploratory findings from semi-structured interviews with ten U.S. cities. *Transport Policy*, 92(C), 1–7. https://doi.org/10.1016/j.tranpol.2020.03.005

Campbell KB, Brakewood C (2017) Sharing riders: How bikesharing impacts bus ridership in New York City. *Transportation Research*. Part A, Policy and Practice 100:264–282. https://doi.org/10.1016/j.tra.2017.04.017

Chen Z, van Lierop D, Ettema D (2020) Dockless bike-sharing systems: what are the implications? Transport Reviews 40(3):333–353. https://doi.org/10.1080/01441647.2019.1710306

Cottrill CD (2020) MaaS surveillance: Privacy considerations in mobility as a service. *Transportation Research*. Part A, Policy and Practice 131:50–57. https://doi.org/10.1016/j.tra.2019.09.026

Fishman E, Washington S, Haworth N (2014) Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia. *Transportation Research*. Part D, Transport and Environment 31:13–20. https://doi.org/10.1016/j.trd.2014.05.013

Frias-Martinez, V., Sloate, E., Manglunia, H., & Wu, J. (2021). Causal effect of low-income areas on shared dockless e-scooter use. *Transportation Research. Part D, Transport and Environment*, 100, 103038. https://doi.org/10.1016/j.trd.2021.103038

Grise, E., & El-Geneidy, A. (2018). Where is the happy transit rider? Evaluating satisfaction with regional rail service using a spatial segmentation approach. *Transportation Research. Part A, Policy and Practice*, 114, 84–96. https://doi.org/10.1016/j.tra.2017.11.005

Hamilton TL, Wichman CJ (2018) Bicycle infrastructure and traffic congestion: Evidence from DC's Capital Bikeshare. Journal of Environmental Economics and Management 87:72–93. https://doi.org/10.1016/j.jeem.2017.03.007

He, P., Zou, Z., Zhang, Y., & Baiocchi, G. (2020). Boosting the eco-friendly sharing economy: the effect of gasoline prices on bikeshare ridership in three u.s. metropolises. *Environmental Research Letters,* 15(11), 114021–114021. https://doi.org/10.1088/1748-9326/abbb52

Herlihy DV (2004) Bicycle : the history. Yale University Press

Hollingsworth J, Copeland B, Johnson JX (2019) Are e-scooters polluters? The environmental impacts of shared dockless electric scooters. Environmental Research Letters 14(8):84031. https://doi.org/10.1088/1748-9326/ab2da8

Huré M, Passalacqua A (2017) La Rochelle, France, and the invention of bike sharing public policy in the 1970s. Journal of Transport History 38(1):106–123. https://doi.org/10.1177/0022526616676275

Hyra, D. (2016). Commentary. *Cityscape* (Washington, D.C.), 18(3), 169–178.

Kelley, T. (2016). Riding Arlington's Capital Bikeshare Stations. https://www.bikearlington.com/riding-all-of-arlingtons-capital-bikeshare-stations/

Kong, H., Jin, S. T., & Sui, D. Z. (2020). Deciphering the relationship between bikesharing and public transit: Modal substitution, integration, and complementation. *Transportation Research. Part D, Transport and Environment*, 85, 102392. https://doi.org/10.1016/j.trd.2020.102392

Lang, N., Schellong, D., Hagenmaier, M., Herrmann, A., & Hohenreuther, M. (2022). Putting Micromobility at the Center of Urban Mobility. https://www.bcg.com/publications/2022/the-future-of-urban-mobility

Leszczynski, A., & Kong, V. (2022). Walking (with) the platform: bikesharing and the aesthetics of gentrification in Vancouver. *Urban Geography*, ahead-of-print(ahead-of-print), 1–23. https://doi.org/10.1080/02723638.2022.2036926

Luo H, Kou Z, Zhao F, Cai H (2019) Comparative life cycle assessment of station-based and dockless bike sharing systems. Resources, Conservation and Recycling 146:180–189. https://doi.org/10.1016/j.resconrec.2019.03.003

Ma T, Liu C, Erdoğan S (2015) Bicycle Sharing and Public Transit. Transportation Research Record 2534(1):1–9. https://doi.org/10.3141/2534-01

McKenzie G (2020) Urban mobility in the sharing economy: A spatiotemporal comparison of shared mobility services. Computers, Environment and Urban Systems 79:101418. https://doi.org/10.1016/j.compenvurbsys.2019.101418

Mitra R, Hess PM (2021) Who are the potential users of shared e-scooters? An examination of socio-demographic, attitudinal and environmental factors. Travel, Behaviour & Society 23:100–107. https://doi.org/10.1016/j.tbs.2020.12.004

Oeschger, G., Carroll, P., & Caulfield, B. (2020). Micromobility and public transport integration: The current state of knowledge. *Transportation Research. Part D, Transport and Environment*, 89, 102628. https://doi.org/10.1016/j.trd.2020.102628

Orozco-Fontalvo, M., Llerena, L., & Cantillo, V. (2022). Dockless electric scooters: A review of a growing micromobility mode. *International Journal of Sustainable Transportation*, ahead-of-print(ahead-of-print), 1–17. https://doi.org/10.1080/15568318.2022.2044097

Parkes SD, Marsden G, Shaheen SA, Cohen AP (2013) Understanding the diffusion of public bikesharing systems: evidence from Europe and North America. Journal of Transport Geography 31:94–103. https://doi.org/10.1016/j.jtrangeo.2013.06.003

Qian X, Jaller M (2020) Bikesharing, equity, and disadvantaged communities: A case study in Chicago. *Transportation Research*. Part A, Policy and Practice 140:354–371. https://doi.org/10.1016/j.tra.2020.07.004

Riggs W, Kawashima M, Batstone D (2021) Exploring best practice for municipal e-scooter policy in the United States. *Transportation Research*. Part A, Policy and Practice 151:18–27. https://doi.org/10.1016/j.tra.2021.06.025

Shaheen SA, Guzman S, Zhang H (2010) Bikesharing in Europe, the Americas, and Asia. Transportation Research Record 2143(1):159–167. https://doi.org/10.3141/2143-20

Stahley, L. H., Camhi, S. M., Wright, J. A., & Troped, P. J. (2022). Correlates of bike share use and its association with weight status at an urban university. *PloS One*, 17(8), e0270870–e0270870. https://doi.org/10.1371/journal.pone.0270870

Taylor, A. The Bike-Share Oversupply in China: Huge Piles of Abandoned and Broken Bicycles. https://www.theatlantic.com/photo/2018/03/bike-share-oversupply-in-china-huge-piles-of-abandoned-and-broken-bicycles/556268/

Teixeira João Filipe, Lopes Miguel (2020) The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York's Citi Bike. Transportation Research Interdisciplinary Perspectives 6:100166

Topos Magazine. (2016). Cykelslangen Copenhagen: An Urban Snake. https://toposmagazine.com/copenhagen-cykelslangen/#Rasmus-Hjortshoj-Cykelslangen-LARGE-16-631%C3%97440

Transportation for America. (2022). Shared Micromobility Playbook. https://playbook.t4america.org/

Triple M. (2021). New E-Scooter Safety Trial Launches In Brisbane. https://www.triplem.com.au/story/new-e-scooter-safety-trial-launches-in-brisbane-189954

The Brisbane City Council. (2022). Brisbane's e-mobility strategy. https://www.brisbane.qld.gov.au/traffic-and-transport/transport-plan-for-brisbane/transport-plan-for-brisbane-implementation-plan/brisbanes-e-mobility-strategy

The North American Bikeshare & Scootershare Association (NABSA). (2021). 3rd Annual Shared Micromobility: State of the Industry Report. https://nabsa.net/about/industry/

The National Association of City Transportation Officials (NACTO). (2018). Bike Share in the U.S.: 2017. https://nacto.org/bike-share-statistics-2017/

The National Association of City Transportation Officials (NACTO). (2020). Shared Micromobility in the U.S.: 2019. https://nacto.org/shared-micromobility-2019/

Van der Zee, R. (2016). Story of cities #30: how this Amsterdam inventor gave bike-sharing to the world. In the Guardian: https://www.theguardian.com/cities/2016/apr/26/story-cities-amsterdam-bike-share-scheme

Wang, K., Qian, X., Fitch, D. T., Lee, Y., Malik, J., & Circella, G. (2022). What travel modes do shared e-scooters displace? A review of recent research findings. *Transport Reviews*, ahead-of-print(ahead-of-print), 1–27. https://doi.org/10.1080/01441647.2021.2015639

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., & Goodman, A. (2014). Health effects of the London bicycle sharing system: health impact modelling study. *BMJ (Online)*, 348(feb13 1), g425–g425. https://doi.org/10.1136/bmj.g425

Xu D (2019) Burn Calories, Not Fuel! The effects of bikeshare programs on obesity rates. *Transportation Research.* Part D, Transport and Environment 67:89–108. https://doi.org/10.1016/j.trd.2018.11.002

Yang Y, Heppenstall A, Turner A, Comber A (2019) A spatiotemporal and graph-based analysis of dockless bike sharing patterns to understand urban flows over the last mile. Computers, Environment and Urban Systems 77:101361. https://doi.org/10.1016/j.compenvurbsys.2019.101361

Zhang Y, Mi Z (2018) Environmental benefits of bike sharing: A big data-based analysis. Applied Energy 220:296–301. https://doi.org/10.1016/j.apenergy.2018.03.101

Zhou X (2015) Understanding spatiotemporal patterns of biking behavior by analyzing massive bike-sharing data in Chicago. PloS One 10(10):e0137922–e0137922. https://doi.org/10.1371/journal.pone.0137922

Zou. Z. (2021). Planning towards an equitable sharing economy: On housing, on transportation, on policymaking. Doctoral dissertation, available at: https://doi.org/10.13016/mnq5-4cq2

Zou, Z., Younes, H., Erdoğan Sevgi, & Wu, J. (2020). Exploratory analysis of real-time e-scooter trip data in Washington, D.C. *Transportation Research Record: Journal of the Transportation Research Board*, 2674(8), 285–299. https://doi.org/10.1177/0361198120919760

Zuniga-Garcia, N., Tec, M., Scott, J. G., & Machemehl, R. B. (2022). Evaluation of e-scooters as transit last-mile solution. Transportation Research. Part C, Emerging Technologies, 139, 103660. https://doi.org/10.1016/j.trc.2022.103660

Chapter 7
Understanding Bikeability: Insight into the Cycling-City Relationship Using Massive Dockless Bike-Sharing Records in Beijing

Enjia Zhang, Wanting Hsu, Ying Long, and Scott Hawken

Abstract Cycling records from emerging dockless bike-sharing services provide new opportunities to gain insight into the interactions between multiple fine-scale cycling characteristics and built environmental elements. Using Beijing as an example and the street as the analytic unit, this study examined the associations between three cycling characteristics and spatial visual elements while controlling for other built environmental features. The results showed that most visual elements were significantly associated with cycling characteristics, but their performance differs across models for trip distance, speed, and volume. The results also indicated that individuals riding long distances or at fast speeds preferred streets with more sky and greenery views. Likewise, wider streets with less spatial disorder, tended to have a higher riding volume. The findings can enhance the understanding of cycling behaviors and promote the implementation of urban design for more bikeable streets.

Keywords Dockless bike-sharing · Bikeability · Cycling characteristics · Spatial visual elements · Beijing

7.1 Introduction

Cycling is believed to promote the sustainable development of cities by providing a low-emission solution for commuting and recreational travel, especially in high-density cities where it can help address the last-mile problem (Nogal and Jimenez 2020). Additionally, it is considered a physical activity that brings health benefits

E. Zhang · W. Hsu · Y. Long (✉)
School of Architecture, Tsinghua University, Beijing, China
e-mail: ylong@tsinghua.edu.cn

E. Zhang
e-mail: zej18@mails.tsinghua.edu.cn

S. Hawken
School of Architecture and Civil Engineering, The University of Adelaide, Adelaide, Australia
e-mail: scott.hawken@adelaide.edu.au

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_7

to individuals (Otero et al. 2018). Therefore, cycling behavior (Kaplan et al. 2015; Castanon & Ribeiro 2021) and influential elements from various aspects (Castanon & Ribeiro 2021; Hardinghaus et al. 2021; Shaer et al. 2021) have long been a topic of interest for scholars in transportation, urban planning, and public health (Forsyth & Krizek 2011; Hu et al. 2021; Li et al. 2021).

Benefiting from the development of information and communication technologies, docked and dockless IT-based bike-sharing, as emerging modes of cycling, have witnessed rapid growth recently (Pons et al. 2016; Chen et al. 2020). Meanwhile, IT-based bike-sharing can collect massive cycling records, enabling quantitative studies of the spatiotemporal behaviors and spatial preferences of cyclists. Previous studies have measured cycling behaviors and uncovered associated built environmental features, such as the density and distance of facilities, bike station attributes, geographic altitude, walkscore, street network, and mixed land use, based on the pick-up and drop-off data from bike stations (Faghih-Imani et al. 2014; El-Assi et al. 2017; Scott and Ciuro 2019).

Compared to bike-sharing based on docking stations, dockless bike-sharing allows users to pick up and drop off bicycles anywhere within a service zone (Orvin & Fatmi 2021). Dockless bike-sharing has the potential to effectively promote active travel, improve user mobility, encourage more users to participate in cycling (Orvin & Fatmi 2021), improve the efficiency of bicycle utilization (Tao & Zhou 2021), and extend the transfer radius of public transportation (Ai et al. 2019), in light of the delivered demand-responsive, multimodal services (Shaheen et al. 2012) and flexible access to public transportation (Duran-Rodas et al. 2020). Moreover, dockless shared bikes can collect more detailed and fine-scale cycling data for each street during a user's ride. Therefore, there has been a surge in research on the characteristics of dockless bike-sharing and its relationship with the built environment (Fan & Zheng 2020; Su et al. 2020; Li et al. 2021).

However, most studies have focused on cyclists' route choices, transfers with other public transportation, and bicycle parking, while failing to measure and compare other cycling characteristics, such as speed and distance. Furthermore, although some objective and perceived built environments, such as the distance to subway/bus stations, mixed land use, and the density of residential and office functions and buildings, have been shown to be highly associated with cycling trips (Scott & Ciuro 2019; Li et al. 2021; Guo & He 2021), the spatial visual elements in the streets (Goodspeed & Yan 2017) that urban designers and governments frequently highlight in urban design guidelines (Tang & Long 2019) have not been considered in these studies.

To address this research gap, this study used Beijing as a study area. The study analyses data from the bike-sharing company Mobike to portray three cycling characteristics, and compared their different relationships with spatial visual elements, while also controlling for other built environmental elements with the potential to influence cycling behaviors.

7.2 Methodology

7.2.1 Research Design

This study focused on the area within Beijing's Fifth Ring Road (667 km^2), which is the main urban built-up area for most commuter trips by all modes of transportation in Beijing. The analytic unit was a street segment, which is the portion of the road between two road intersections. Ordinary least squares (OLS) regression was used to examine the relationship between cycling characteristics and spatial visual elements. Meanwhile, we controlled for other built environmental factors that may influence travel demand (Ewing & Cervero 2010). Three cycling characteristics were examined in this study: average trip distance, average trip speed, and daily trip volume. Prior to regression analysis, data distribution was checked, and multicollinearity between independent variables and control variables was assessed (Fig. 7.1).

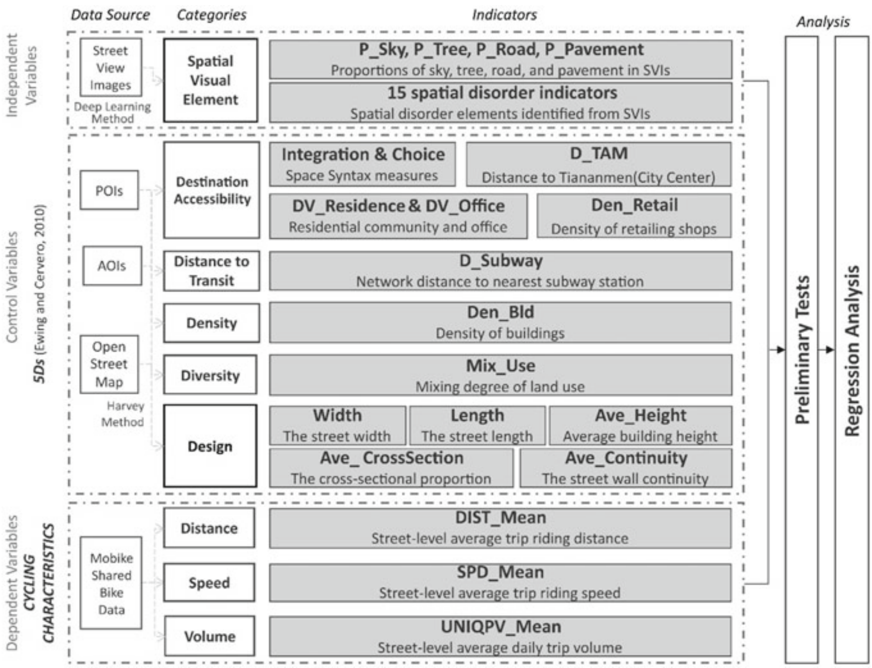


Fig. 7.1 Framework of this study

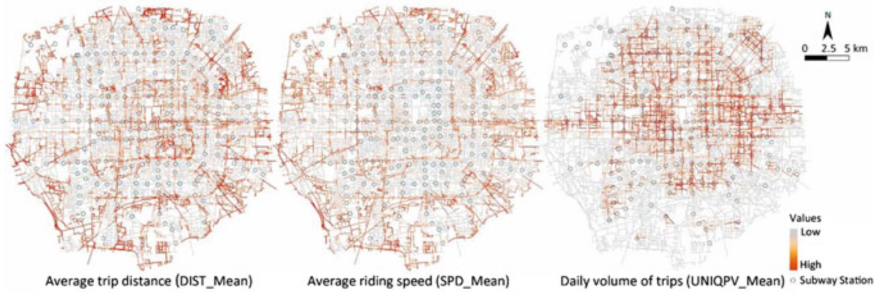


Fig. 7.2 Spatial visualization of three cycling characteristics using Jenks natural breaks

7.2.2 Variables and Data

7.2.2.1 Dependent Variable: Measurement of Cycling Characteristics

Dockless bike-sharing data were collected from Mobike, which was established in China in 2015 and quickly became one of the most popular bike-sharing companies. It was acquired by the e-commerce giant Meituan in April 2018. We collected anonymous bike trip records for the area within Beijing's Fifth Ring Road over a six-month period (181 days from January 1 to June 30, 2018), which were aggregated by street segment. The data included the street ID, date, daily user volume, daily bike trip volume, user's average speed, and average trip distance. Three indicators were calculated and used in this study: the average trip distance of all users who passed through a street segment (DIST_Mean), the average riding speed of all users who passed through a street segment (SPD_Mean), and the daily volume of Mobike trips on each street (UNIQPV_Mean), which were used to depict the distance, speed, and volume of cycling for each street. The data were anonymized to protect user privacy.

Figure 7.2 presents a spatial visualization of the three dependent variables. The average trip distances on the outer-ring streets were higher than those on the fourth-ring road and lower near subway stations. The spatial distribution of the average speed shared some commonalities with the previous map of average trip distances. The trip volume map showed that Mobike trips were concentrated in the city center and near subway stations. The differentiated patterns of these indicators indicate different associated built environmental elements.

7.2.2.2 Independent Variable: Measurement of Spatial Visual Elements

Spatial Visual Elements. The spatial visual elements considered as independent variables in this study were elements viewed from the street, which could be extracted from street-view images (SVI). We obtained the SVIs by crawling Tencent Map using its application programming interface (API). For all streets processed within the fifth ring road area, we divided each street segment into vertices with a distance of 50 m, resulting in an average of four points for each street to collect SVIs that depict the overall visual conditions of each street regarding the continuity of street elements and landscape. The corresponding vertex coordinates were inputted into the place ID retrieval API and the API for downloading the SVIs of four horizontal angles: front, back, left, and right. As a result, several vertices spaced 50 m apart were distributed along the street for each street, providing us with 4-direction scenes for each vertex that may represent the overall spatial elements of a specific street.

Based on the SVIs, we used the SegNet pixel-wise image semantic segmentation method (Badrinarayanan et al. 2017) to calculate the proportions of the sky (P_Sky), trees (P_Tree), road (P_Road), and pavement (P_Pavement) to depict the streets' beauty (greenery and openness) and convivence for riding (road and pavement). Additionally, we measured 15 disorder indicators that could influence the perceived safety for cyclists (Kyttä et al. 2014) by applying the deep learning model proposed by Chen et al. (2022). Specifically, the 15 disorder indicators were abandoned buildings (Bld_Abandoned), buildings with damaged facades (BldFac_Damaged), buildings with unkempt facades (BldFac_Unkempt), graffiti/illegal advertisements (Adver_Graffiti), illegal/temporary buildings (Bld_Illegal), stores with poor signboards (Store_Poorsign), stores with poor facades (Store_Poorfac), vacant and pending stores (Store_Vacant), messy and unmaintained greening (Unmaingreen_Messy), garbage/litter on the street (Garbage), construction fence remnants (Fence_Remnant), broken roads (Road_Broken), occupied roads (Road_Occupied), broken infrastructure (Infra_Broken), and damaged public interfaces (Interface_Damaged). For each observed SVI point, we scored the presence of the above elements as 1; otherwise, it was 0. Thus, for each street, the average score for each disorder variable reflected the average degree of disorder.

7.2.2.3 Control Variable: Measurement of Five Ds

The five Ds (Destination Accessibility, Distance to Transit, Density, Diversity, and Design) have been identified as influential built environments that can moderate travel demands (Ewing & Cervero 2010). Therefore, this study introduced five Ds as control variables to better reveal the relationships between spatial visual indicators and cycling characteristics. Table 7.1 shows the descriptions of all variables.

Destination Accessibility. As the origin–destination is the primary determinant of the cycling route, two space syntax indicators were utilized to control the role of the street network in influencing the route preferences of cyclists: (1) the integration index, which gauges a street segment's ability to attract incoming traffic and reflects its centrality within the entire system, and (2) the choice index, which evaluates the benefits of a spatial unit as the shortest travel path and reflects the possibility of a street segment being traversed (Hillier 1999). We computed the integration and choice measures using analysis radii of 800 m (Int800, Cho800), 1600 m (Int1600, Cho1600), 2400 m (Int2400, Cho2400), 3200 m (Int3200, Cho3200), 4800 m (Int4800, Cho4800), 9600 m (Int9600, Cho9600), and n (global analysis, Intall, Choall), respectively. To evaluate access to the city center, we calculated the

Table 7.1 Descriptions of all variables (N = 16,266)

Variables	Indicator categories	Variables	Descriptions and methods	Mean	Std_Deviation	Unit
Dependent variables	Cycling characteristics	DIST_Mean	The average trip distance of all the users that passed through a street segment	4.3869	1.4587	km
		SPD_Mean	The average riding speed of all the users that passed through a street segment	10.2943	1.1337	km/h
		UNIQPV_Mean	The daily volume of Mobike trips on each street	173.4124	181.4656	#
Independent variables	Spatial visual elements	P_Sky	The average street-level proportions of the sky, trees, road, and pavement in all SVIs within the same street	0.1633	0.0930	–
		P_Tree		0.1948	0.1287	–
		P_Road		0.1540	0.0691	–
		P_Pavement		0.0843	0.0478	–
		Bld_Abandoned	The average score of identified spatial disorder elements in SVIs using MobileNet V3-small based on the results of Chen et al.'s (2022) within the same street	0.0042	0.0246	–
		BldFac_Damaged		0.0699	0.1163	–
		BldFac_Unkempt		0.0486	0.1037	–
		Adver_Graffiti		0.3661	0.2649	–
		Bld_Illegal		0.0097	0.0366	–
		Store_Poorsign		0.4853	0.2390	–
		Store_Poorfac		0.0536	0.0885	–
		Store_Vacant		0.0377	0.0751	–
		Unmaingreen_Messy		0.0179	0.0624	–
		Garbage		0.3250	0.2368	–
		Fence_Remnant		0.0840	0.1386	–

(continued)

Table 7.1 (continued)

Variables	Indicator categories	Variables	Descriptions and methods	Mean	Std_Deviation	Unit
		Road_Broken		0.2535	0.2366	–
		Road_Occupied		0.0492	0.1015	–
		Infra_Broken		0.1295	0.1922	–
		Interface_Damaged		0.1176	0.1706	–
Control Variables	Destination accessibility	Integration (Int800, Int1600, Int2400, Int3200, Int4800, Int9600, Intall)	Integration index value using space syntax analysis with analysis radius of 800 m, 1600 m, 2400 m, 3200 m, 4800 m, 9600 m, and n (global)	–	–	–
		Choice (Cho800, Cho1600, Cho2400, Cho3200, Cho4800, Cho9600, Choall)	Choice index value using space syntax analysis with analysis radius of 800 m, 1600 m, 2400 m, 3200 m, 4800 m, 9600 m, and n (global)	–	–	–
		D_TAM	The Euclidean distance to Tiananmen Square	8.6852	3.6139	km
		DV_Residence	Dummy variable: when the street is near residential communities (within 100 m), the score is 1, otherwise 0	0.7200	0.4500	–
		DV_Office	Dummy variable: when the street is near offices (within 100 m), the score is 1, otherwise 0	0.4000	0.4890	–

(continued)

Table 7.1 (continued)

Variables	Indicator categories	Variables	Descriptions and methods	Mean	Std_Deviation	Unit
		Den_Retail	The density of POIs of retail stores within each street's 50-m buffer	375.7777	636.824	#/km^2
	Distance to transit	D_Subway	The network distance to the nearest subway station	0.9811	0.8598	km
	Density	Den_Bld	The average number of buildings on two sides of a street	22.9867	16.8490	#/km^2
	Diversity	Mix_Use	The mixing degree of POIs using Shannon's entropy	0.5344	0.1998	–
	Design	Width	Distance between edges across the street	35.3800	18.7200	m
		Length	Centerline distance along a street	248.6848	179.9437	m
		Ave_Height	Average building height along the street	11.8000	14.5330	m
		Ave_CrossSection	Width/average height on both sides of the street	0.3766	0.5571	–
		Ave_Continuity	The average proportion of edge intersecting buildings on both sides of the street	0.3345	0.2325	–

Note We collected road networks, building footprints, AOI, and POI from Gaode map (https://lbs.amap.com) from its open API distance from the midpoint of each street to the flag point base in Tiananmen Square (D_TAM). We also considered several crucial sites that could be potential origins or destinations for the rides, including dummy variables for residential communities (DV_Residence) and offices (DV_Office) within a 100-m distance to the street in the Area of Interest (AOI) data, and the density of retail stores (shopping and catering) within a 50-m buffer of the street (Den_Retail).

Distance to Transit. The network distance from the street segment to the nearest subway station (D_Subway) was regarded as the distance to transit.

Density. Density indicators were measured as building counts divided by the street length (Den_Bld).

Diversity. Diversity measures the mixing degree of land use in 50-buffer streets (Mix-Use). The normalized proportion of each main category of Point of Interest (POI) was calculated using Shannon's entropy.

Design. Some street-level urban forms in the design category, such as the width (Width) and length (Length) of the street, average height (Ave_Height) and continuity (Ave_Continuity) of surrounding buildings, and the average cross-section (street width/building height) (Ave_CrossSection) were also calculated by referring to the GIS-based methods developed by Harvey (2014).

7.3 Results

7.3.1 Data Processing and Preliminary Tests

Before conducting the regression analysis, we checked the distributions of all variables. Since Den_Retail and UNIQPV_Mean were long-tailed data, we used the log transformation on these two variables to ensure the reliability of the models. Then, we applied Pearson's correlation and variance inflation factor (VIF) tests to avoid the multicollinearity effect. The results showed that the multicollinearity of the model was not severe, with Pearson's correlation coefficients less than 0.8, and VIF values less than 5.

To ensure that the OLS model performed better, we calculated the Pearson correlations between the various space syntax measures and the three cycling characteristic indicators. We selected those with higher coefficients to be used in the following regression analysis. The results showed that for DIST_Mean, Int800 and ChoAll had the highest values; for SPD_Mean, Int1600 and Cho800 showed a closer relationship; and for LnUNIQPV_Mean, Int3200 and Cho3200 presented the highest coefficients. Therefore, this study considered different integration and choice variables in the three regression models.

7.3.2 Regression Analysis and Results

Table 7.2 displays the results of the regression models with different dependent variables. The results showed that the built environmental variables could explain 41.2% of the trip distance, 34.8% of the cycling speed, and 54.9% of the trip volume, suggesting that trip volume has a stronger relationship with built environmental elements than trip speed and distance.

Among all the spatial visual variables, the proportions of roads and pavements in the SVIs were significantly associated with all three cycling characteristics, while most indicators such as P_Sky, P_Tree, Bld_Abandoned, BldFac_Unkempt, Adver_Graffiti, Bld_Illegal, Store_Poorsign, Store_Poorfac, Unmaingreen_Messy, Garbage, Road_Broken, and Road_Occupied, were only relevant to cycling characteristics in specific contexts.

The results revealed that people preferred to ride at a faster speed on streets with broader views of sky, greenery, roads, and pavements, as all the proportions of sky, trees, road, and pavement in the SVIs were significantly positive with trip distance and cycling speed. For trip volume, the proportion of roads had a positive correlation, but the pavement proportion had adverse effects. This implies that wider roads with narrower sidewalks were more likely to witness more bike-sharing trips.

The findings for the spatial disorder indicators showed that different cycling characteristics were related to diverse elements. Long-distance rides usually occurred in areas with poor spatial quality, such as abandoned buildings, unkempt building façades, poor store façades, and broken roads. This implies that people who lived near urban villages (usually with poor spatial quality) tended to use shared bikes for long-distance commuting.

The results for speed suggested that people would quickly pass places with garbage and slow down along streets with unkempt illegal/temporary buildings and messy and unmaintained greening. One possible explanation is that places with poor greenery and temporary buildings are usually commercial or residential in suburban areas, which could be destinations for riders.

As for trip volumes, some small elements, such as Adver_Graffiti, Store_Poorsign, and Garbage, had positive coefficients, whereas some larger items, such as unkempt buildings, messy and unmaintained greenery and busy car-filled roads, were negatively associated. This implies that disorder in buildings, landscapes, and roads could hinder people's path choices for cycling.

Figure 7.3 presents the standardized coefficients for the significant indicators, allowing for a more direct interpretation of the differences in the regression results among different cycling characteristics. The results showed that street network features, potential origin and destination places, access to subway stations, mixed land use, and some urban form indicators were highly associated with cycling characteristics, which is consistent with previous studies (Guo & He 2021; Zhuang et al. 2022). The results also suggest that visual elements on the street, such as the proportion of sky, trees, roads, and pavement, are much more critical for riders than spatial disorder indicators on the two sides of the street. Moreover, more spatial disorder

Table 7.2 Comparison between results using different cycling behaviors as dependent variables

N = 16,266	Variable description	DIST_Mean	SPD_Mean	LnUNIQPV_Mean
Destination accessibility	Int800	−1.446[c] (0.421)		
	Int1600		−1.198[c] (0.190)	
	Int3200			2.097[c] (0.090)
	Cho800		−0.036[b] (0.014)	
	Cho3200			0.000 (0.000)
	ChoAll	0.000[c] (0.000)		
	D_TAM	−0.082[c] (0.003)	0.011[c] (0.003)	−0.011[c] (0.003)
	DV_Residential	−0.501[c] (0.019)	−0.319[c] (0.019)	0.517[c] (0.019)
	DV_Office	0.048[b] (0.017)	0.202[c] (0.016)	0.168[c] (0.017)
	LnDen_Retail	0.008 (0.004)	−0.011[a] (0.004)	0.051[c] (0.005)
Distance to transit	D_Subway	0.400[c] (0.011)	0.171[c] (0.011)	−0.316[c] (0.011)
Density	Den_Bld	−0.004[c] (0.001)	−0.007[c] (0.001)	−0.007[c] (0.001)
Diversity	Mix_Use	−1.248[c] (0.065)	−1.195[c] (0.063)	1.820[c] (0.066)
Design	Width	−0.002[c] (0.000)	0.000 (0.000)	0.007[c] (0.000)
	Length	0.000[c] (0.000)	0.001[c] (0.000)	0.000[c] (0.000)
	Ave_Height	−0.002[a] (0.001)	0.001 (0.001)	−0.001[a] (0.001)
	Ave_Section	−0.100[c] (0.017)	−0.067[c] (0.016)	−0.057[c] (0.017)
	Ave_Continuity	−0.278[c] (0.053)	−0.154[a] (0.051)	−0.346[c] (0.053)
Spatial visual element	P_Sky	3.375c (0.136)	1.819[c] (0.130)	−0.005 (0.136)
	P_Tree	1.136[c] (0.074)	0.416[c] (0.071)	−0.144 (0.074)
	P_Road	1.072[c] (0.156)	0.867[c] (0.149)	2.871[c] (0.157)

(continued)

Table 7.2 (continued)

N = 16,266	Variable description	DIST_Mean	SPD_Mean	LnUNIQPV_Mean
	P_Pavement	1.845[c] (0.177)	0.490[b] (0.170)	−1.296[c] (0.178)
	Bld_Abandoned	0.646[a] (0.315)	0.290 (0.301)	−0.055 (0.317)
	BldFac_Damaged	−0.105 (0.077)	0.008 (0.074)	−0.114 (0.078)
	BldFac_Unkempt	0.420[c] (0.088)	0.094 (0.084)	−0.650[c] (0.089)
	Adver_Graffiti	−0.049 (0.029)	−0.043 (0.027)	0.192[c] (0.029)
	Bld_Illegal	−0.392 (0.216)	−0.548[b] (0.207)	0.104 (0.217)
	Store_Poorsign	−0.052 (0.031)	−0.016 (0.029)	0.161[c] (0.031)
	Store_Poorfac	0.342[c] (0.096)	−0.073 (0.092)	−0.176 (0.097)
	Store_Vacant	−0.189 (0.110)	−0.119 (0.105)	−0.006 (0.110)
	Unmaingreen_Messy	−0.185 (0.126)	−0.248[a] (0.120)	−0.598[c] (0.127)
	Garbage	0.059 (0.032)	0.068[a] (0.031)	0.107[c] (0.032)
	Fence_Remnant	0.051 (0.058)	−0.082 (0.055)	0.028 (0.058)
	Road_Broken	0.070[a] (0.032)	0.044 (0.031)	0.001 (0.032)
	Road_Occupied	−0.057 (0.082)	−0.042 (0.078)	−0.223[b] (0.082)
	Infra_Broken	−0.022 (0.046)	0.077 (0.044)	−0.013 (0.047)
	Interface_Damaged	0.034 (0.044)	−0.017 (0.043)	0.016 (0.045)
R^2		0.414	0.349	0.550
Adjusted R^2		0.412	0.348	0.549

Note The table reports the coefficients and predictive power (R^2) for each model's column. Standard errors are in parentheses. Significance level: [a]$p < 0.05$, [b]$p < 0.01$, and [c]$p < 0.001$

elements were associated with trip volume than with distance and speed, while more visual proportion indicators from the street view were significantly correlated with trip distance and speed than with volume. These results reflected riders' varying preferences for spatial visual elements.

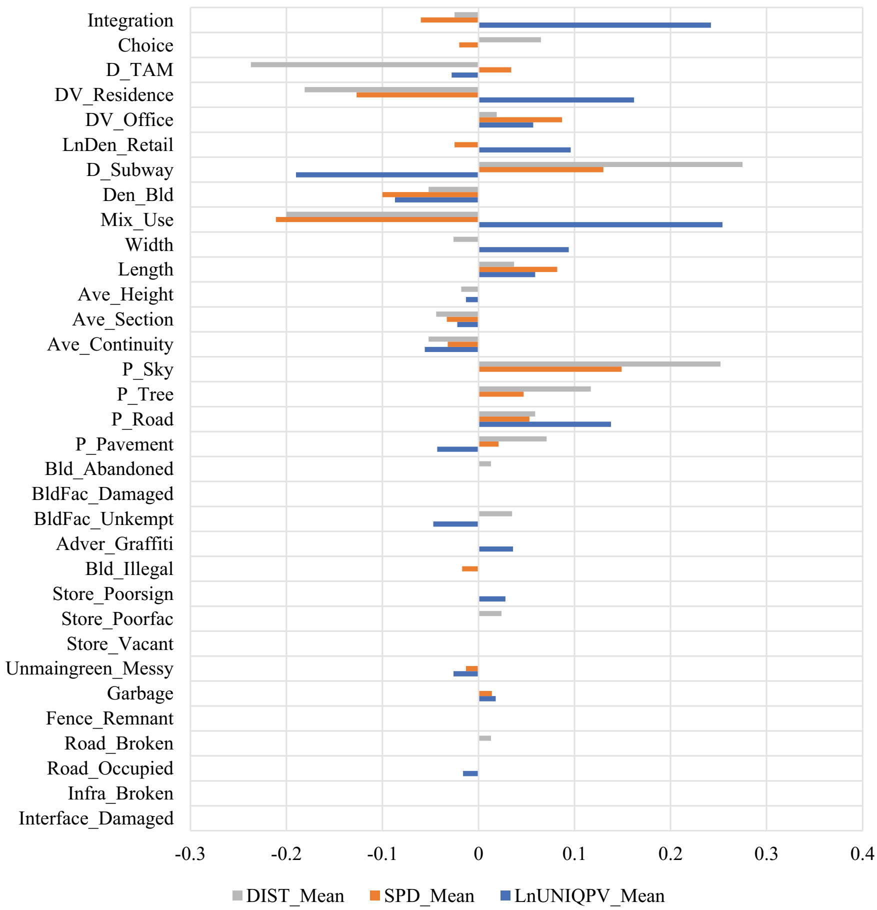


Fig. 7.3 Comparison between the standardized coefficients of statistically significant indicators ($p < 0.05$) for three cycling characteristics

7.4 Conclusions and Discussion

Taking Beijing's central city as the study area and street segments as the analytic unit, this study examined the relationship between three cycling characteristics (average trip distance, average riding speed, and daily volume) and spatial visual elements by controlling for other potentially influential built environmental factors. After controlling for the role of street segments in the whole road network and the width of the road, the results revealed that individuals who ride long distances or at high speeds preferred a broad vision of sky and trees on the streets, while wider streets with fewer spatial disorder elements in terms of larger items such as buildings, landscapes, and roads could have a higher riding volume. Moreover, the results suggested that the visual proportions of elements on the road were much more critical for riders than spatial disorder indicators on both sides of the street.

The findings could enhance our understanding of cyclists' spatial preferences regarding different riding scenarios (e.g., long-distance riding, high-speed riding, most frequent riding) and promote urban design implementations for more bikeable streets. The results revealed distinct spatial visual elements for different cycling characteristics, which can help urban designers develop corresponding guidelines and encourage people to promote this sustainable commuting mode and improve the overall usage efficiency of the bike-sharing system.

However, there are still some limitations to be addressed in future research. First, due to the limitations of data acquisition, we could only obtain the data aggregated to the street segment, and not each ride's specific OD and path data. With more data, future studies could consider the route choice and direction of cycling to analyze the built environmental elements along the entire ride trip and on the closed side of the street. Second, further studies should expand the data sources to measure bike infrastructure, such as bike lanes and parking areas, to better investigate preferences for cycling and parking.

Funding This work was supported by the National Natural Science Foundation of China grant number 52178044 and the UNSW-TSINGHUA UNIVERSITY Collaborative Research Fund RG180121.

References

Ai Y, Li ZP, Gan M (2019) A solution to measure traveler's transfer tolerance for walking mode and dockless bike-sharing mode. J Supercomput 75(6):3140–3157

Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 39(12):2481–2495

Castanon UN, Ribeiro PJG (2021) Bikeability and emerging phenomena in cycling: exploratory analysis and review. Sustainability 13(4):2394

Chen ZY, van Lierop D, Ettema D (2020) Dockless bike-sharing systems: what are the implications? Transp Rev 40(3):333–353

Chen J, Chen L, Li Y et al (2022) Measuring physical disorder in urban street spaces: a large-scale analysis using street view images and deep learning. Ann Am Assoc Geogr https://doi.org/10.1080/24694452.2022.2114417

Duran-Rodas D, Villeneuve D, Wulfhorst G (2020) Bike-sharing: the good, the bad, and the future -an analysis of the public discussion on twitter. Eur J Transp Infrastruct Res 20(4):38–58

El-Assi W, Mahmoud MS, Habib KN (2017) Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. Transportation 44(3):589–613

Ewing R, Cervero R (2010) Travel and the built environment. J Am Plann Assoc 76(3):265–294

Faghih-Imani A, Eluru N, El-Geneidy AM et al (2014) How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal. J Transp Geogr 41:306–314

Fan YC, Zheng SQ (2020) Dockless bike sharing alleviates road congestion by complementing subway travel: evidence from Beijing. Cities 107:102895

Forsyth A, Krizek K (2011) Urban design: is there a distinctive view from the bicycle? J Urban Des 16(4):531–549

Goodspeed R, Yan X (2017) Crowdsourcing street beauty: Visual preference surveys in the big data era. In: Schintler LA, Chen Z (eds) Big data for regional science. Routledge, London and New York, pp 75–93

Guo Y, He S (2021) The role of objective and perceived built environments in affecting dockless bike-sharing as a feeder mode choice of metro commuting. Transp Res Part A Policy Pract 149:377–396

Hardinghaus M, Nieland S, Lehne M et al (2021) More than bike lanes-a multifactorial index of urban bikeability. Sustainability 13(21):11584

Harvey CW (2014) Measuring streetscape design for livability using spatial data and methods. University of Vermont, Vermont

Hillier B (1999) Space is the machine: a configurational theory of architecture. Cambridge University Press, Cambridge

Hu SH, Xiong CF, Liu ZQ et al (2021) Examining spatiotemporal changing patterns of bike-sharing usage during Covid-19 pandemic. J Transp Geogr 91:102997

Kaplan S, Manca F, Nielsen TAS et al (2015) Intentions to use bike-sharing for holiday cycling: an application of the theory of planned behavior. Tour Manag 47:34–46

Kyttä M, Kuoppa J, Hirvonen J et al (2014) Perceived safety of the retrofit neighborhood: a location-based approach. Urban Des Int 19(4):311–328

Li HW, Xing YY, Zhang WB et al (2021) Investigating the impact of weather conditions and land use on dockless bike-share trips in Shanghai, China. J Urban Plann Dev 147(4):237688371

Nogal M, Jimenez P (2020) Attractiveness of bike-sharing stations from a multi-modal perspective: the role of objective and subjective features. Sustainability 12(21):9062

Orvin MM, Fatmi MR (2021) Why individuals choose dockless bike sharing services? Travel Behav Soc 22:199–206

Otero I, Nieuwenhuijsen MJ, Rojas-Rueda D (2018) Health impacts of bike sharing systems in Europe. Environ Int 115:387–394

Pons, JMS Llado JM, Perez MR et al (2016) Bike-sharing schemes and sustainable urban mobility. An analysis in the city of Palma (Mallorca, Balearic Islands). Boletin de la Asoc de Geogr Espanoles 71:227−245

Scott DM, Ciuro C (2019) What factors influence bike share ridership? an investigation of hamilton, Ontario's bike share hubs. Travel Behav Soc 16:50–58

Shaer A, Rezaei M, Rahimi BM et al (2021) Examining the associations between perceived built environment and active travel, before and after the covid-19 outbreak in Shiraz city. Iran Cities 115:103255

Shaheen S, Guzman S, Zhang H (2012) Bikesharing across the globe. In: City cycling. MIT Press, UC Berkeley: Transportation Sustainability Research Center, p 183.

Su D, Wang YC, Yang N et al (2020) Promoting considerate parking behavior in dockless bike-sharing: an experimental study. Transp Res Part A Policy Pract 140:153–165

Tang J, Long Y (2019) Measuring visual quality of street space and its temporal variation: methodology and its application in the hutong area in Beijing. Landsc Urban Plan 191:103436

Tao J, Zhou ZH (2021) Evaluation of potential contribution of dockless bike-sharing service to sustainable and efficient urban mobility in China. Sustain Prod Consum 27:921–932

Zhuang C, Li S, Tan Z et al (2022) Nonlinear and threshold effects of traffic condition and built environment on dockless bike sharing at street level. J Transp Geogr 102:103375

Chapter 8
Disclosing the Impact of Micro-level Environmental Characteristics on Dockless Bikeshare Trip Volume: A Case Study of Ithaca

Qiwei Song, Wenjing Li, Jintai Li, Xinran Wei, and Waishan Qiu

Abstract Although prior literature has examined the impact of the built environment on cycling behavior, the focus has been confined to macro-level environmental characteristics or limited objective features. The role of perceived qualities measured from visual surveys is largely unknown. Using a large amount of dockless bikeshare trajectories, this study maps the cycling trip volume at the street segment level. The research evaluates the micro-level objective features and perceived qualities along the cycling routes using street view imagery, computer vision, and machine learning. Through several regression models, the strengths of both micro-level environment characteristic groups are comprehensively analyzed to reveal their impacts on cycling volume at the street level. Overall, objective features exhibit higher predictive power than perceived qualities, while perceived qualities can complement objective features. The research justifies the significant impacts of micro-level environment characteristics on cycling route choices. It provides a valuable reference for urban planning toward a sustainable cycling-friendly city.

Keywords Dockless bikeshare · Street view · Perceived qualities · Machine learning · Computer vision

Q. Song
John H. Daniels Faculty of Architecture, Landscape and Design, University of Toronto, Toronto, ON M5S 2J5, Canada
e-mail: qiwei.song@mail.utoronto.ca

W. Li
Center for Spatial Information Science, The University of Tokyo, Chiba 277-8568, Japan
e-mail: liwenjing@csis.u-tokyo.ac.jp

J. Li
Westcliff University, Irvine, CA 92614, USA

X. Wei
College of Architecture and Urban Planning, Tongji University, Shanghai 200092, China
e-mail: lindaweixr@foxmail.com

W. Qiu (✉)
Department of City and Regional Planning, Cornell University, Ithaca, NY 14850, USA
e-mail: qiuwaishan@126.com

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_8

8.1 Introduction

Bikeshare is an indispensable cycling mode that promotes healthy cities, facilitates social wellbeing, encourages sustainable mobility, and increases neighborhood vitality (Chen et al. 2020b; Guo and He 2020; Tarpin-Pitre and Morency 2020). More recently, the dockless bikeshare (DBS) system has been increasingly adopted worldwide because of its various advantages (e.g., accessibility and convenience) over the conventional docked bikeshare system (Gu et al. 2019). Thus, understanding what underlying built environment characteristics can affect DBS cycling behavioral patterns is critical in promoting successful evidence-based city planning.

Among many built environment factors, prior research has provided evidence that macro-level characteristics such as land use types and points of interest (POIs) could inform cycling mobility patterns (Guo and He 2020; Qiu and Chang 2021). More recently, how micro-level characteristics (i.e., human perceived environment attributes at the eye level) can influence cycling behaviors have attracted increasing research interests (Ma and Cao 2019; Guo and He 2021). More specifically, an explicit knowledge of how people engage with the built environment and people's interactions with route choices can allow policymakers to adjust urban planning and design guidelines to support the enduring increase in cycling activities, eventually invigorating urban life through active transportation (Goodman et al. 2014). Nevertheless, several research gaps still need to be addressed further.

First, GPS route trajectories have been used previously to reflect real interactions between cyclists and the built environment. Research has primarily focused on the aggregated impacts using an area as the spatial unit due to data availability (Mertens et al. 2016; Guo and He 2021). However, the impact at the street segment level is less discussed (Aultman-Hall et al. 1997; Qiu and Chang 2021). Furthermore, studies utilizing GPS data only contained limited trajectories, making it challenging to capture the citywide activities and traveling patterns (Winters et al. 2010; Krenn et al. 2014). Thus, DBS cycling route choices as cyclists' revealed preferences in the North American small city context should be mapped at the street level using expanded GPS data to provide more insights (Chen et al. 2020b).

Second, studies that have investigated the impact of the micro-level environment on cycling behavior have primarily concentrated on objective attributes, including perceived land use type, building density, and infrastructure (Kerr et al. 2016; Ma and Dill 2017). However, few researchers have analyzed the micro-level perceived environment characteristics through Street View Imagery (SVI) to reveal their impacts on cycling behavior (Lu et al. 2019; Wang et al. 2020). Using this method, previous research has analyzed two types of the micro-level environment characteristics, which are objective features (i.e., visually experienced street features) and perceived qualities (i.e., overall subjective perceptions) (Zhang et al. 2018; Qiu et al. 2022). Nevertheless, the existing related literature has only examined limited types of objective features (mostly street greenery), while they have completely ignored the perceived qualities (Qiu et al. 2022). Perceived qualities of the cyclists when traveling along the routes can be measured using SVIs, visual survey, Computer Vision (CV), and

Machine Learning (ML) algorithms using a high-throughput framework (Zhang et al. 2018; Song et al. 2022b). The strengths of perceived qualities versus the objective features regarding their influences on trip volume should be better addressed.

Lastly, how macro-level built environment characteristics such as POIs affect route choices at the street segment level should be further revealed (Mateo-Babiano et al. 2016; Guo and He 2020). Specifically, less-discussed road network topological features (e.g., connectivity) described by space syntax theory should be examined (Raford et al. 2007; Law et al. 2014). Therefore, there is value in using expanded GPS cycling route data to understand how macro-level characteristics influence behavioral choices compared to the micro-level environment characteristics.

In summary, leveraging an unprecedented amount of GPS data collected from the DBS system, this study sets out to reflect the desirability of cycling route choices by quantifying trip volume (i.e., trip counts) at the street segment level. The study further investigates whether and to what extent micro-level environment characteristics and macro-level environment characteristics influence cycling route choices. Furthermore within the micro-level environment characteristics, how objective features and perceived qualities complement or conflict with each other in affecting trip volume is examined in detail.

8.2 Literature Review

8.2.1 Perceived Built Environment and Cycling Behavior

The cycling behavior has been reported to be associated with perceived built environmental characteristics, including perceived land use type (Van Dyck et al. 2012), safety (Mertens et al. 2016), and other factors. Guo and He (2021) proposed six perceived categories (i.e., accessibility, availability, attractiveness, supportability, safety, substitutability) to study the association with the DBS volumes. Nevertheless, these studies have many limitations. On the one hand, existing studies evaluate micro-level perceived characteristics using low-throughput methods such as manual ratings and questionnaires, which is labor-intensive, time-consuming, and with a limited sample size of survey respondents. On the other hand, individual bike trajectory data is also limited; how the perceived environment affects citywide cycling behavior at the street segment level needs to be addressed.

More recently, SVI data has enabled efficient and remote audits of the street environment. It can provide ground truth for human-centric and micro-level environment settings on a sizeable geospatial scale (Goodspeed and Yan 2017; Zhang et al. 2018). Several research investigated how objective greenery extracted from SVIs affects cycling behavior, while they focused on its aggregated impacts in a zone rather than at individual street level (Lu et al. 2019; Wang et al. 2020). For instance, Lu et al. (2019) pointed out that the odds of cycling are positively related to street greenery, while the study neglected the impacts of other visual elements. Therefore, objective features should be further expanded to be comprehensive enough to capture people's complex perceptions. In addition, perceived qualities should also be investigated as an alternative type of micro-level characteristics compared to objective features. Their influences on cycling route choices must be better understood using a large GPS dataset (Zhang et al. 2018). Although we have seen a growing body of literature on examining the impact of the micro-level perceived environment using perceived qualities on socioeconomic aspects such as housing prices (Kang et al. 2021; Qiu et al. 2022), more knowledge is needed to empirically comprehend the relationship between cycling routes and overall perceived qualities (i.e., subjectively measured) using SVI data.

8.2.2 SVI, CV, and ML for Micro-level Environment Characteristics

On the one hand, the objective street features can be measured effectively using SVIs through CV algorithms and Deep Learning (DL) methods. Scholars have proposed the eye-level greenery, sky view factor, and building view factor to quantify features such as trees and sky that people view at the eye level (Li et al. 2015; Gong et al. 2018). And by recombining view indices following complex mathematical formulas based on urban design theories, scholars have been able to calculate several design qualities, such as enclosure (Song et al. 2022b; Qiu et al. 2023; Su et al. 2023; Wang et al. 2023).

On the other hand, Lin and Moudon (2010) pointed out that perceived qualities can comprehensively explain human behaviors because they capture the overall experiential qualities. To overcome the limitations (e.g., time and cost-consuming and labor-intensive) of early research that collected perceived qualities of the built environment through visual surveys (Ewing and Handy 2009), scholars have proposed to collect opinions from crowdsourcing online platforms or a panel of experts, and predict the perceptual qualities mapping in larger geographical regions through integration with CV and ML algorithms (Dubey et al. 2016; Goodspeed and Yan 2017; Zhang et al. 2018). More recent urban planning literature has applied this method to evaluate perceived urban design qualities (Tian et al. 2021; Qiu et al. 2022, 2023; Song et al. 2022a).

8.3 Dataset and Methodology

8.3.1 Study Area

Located in Tompkins County of upstate New York, the City of Ithaca covers 5.39 mile^2 land (in 2010). The city inhabits more than 30,000 residents, with around

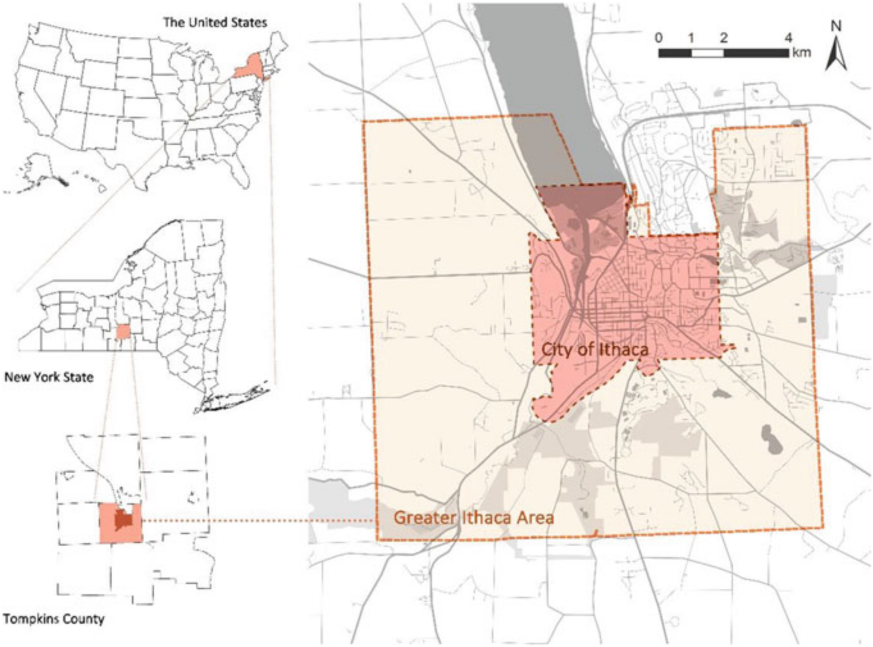


Fig. 8.1 Study area

20,000 students from Cornell University, who make up roughly 2/3 of the population. 49.9% of people are female, while 68.4% are white. Our study area covers the entire service zone (roughly 22.5 mile^2) of dockless bikeshare operation provided by the transportation company 'Lime' (https://www.li.me/en-co) in and around the city (Fig. 8.1).

8.3.2 Methods

First, we calculate the DBS trip volume of each road segment from a dataset that contains massive GPS trajectory records. Second, we quantified the micro-level environment characteristics (including objective features and perceived qualities) using SVI, CV, and ML at individual street level. Third, we collect and compute three types of macro-level characteristics (POI attributes, accessibility attributes, and sociodemographic attributes) that may influence the DBS trip volume in the City of Ithaca. Fourth, using the data above, we build several Ordinary Least Squares (OLS) regression models to examine the correlations and strength of each variable for further detailed comparison.

8.3.3 Data Collection and Processing

8.3.3.1 Dependent Variable: Bikeshare Trip Volume

Raw Lime dockless bikeshare trip trajectories: The GPS data was provided by the company Lime. The City of Ithaca partnered exclusively with Lime to explore bikeshare as an active transportation solution in the fall of 2017. In 2018, Lime launched 350 dockless bikes in Ithaca. The dataset was collected from mobile phones with enabled Auto-GPS function through the installed Lime app. The dataset does not include specific user information such as name, age, or gender for data privacy reasons as part of the agreed terms with users. For each trip of the DBS system, the data contains a unique trip ID, GPS coordinates (latitude and longitude), and timestamps with a 1-s interval. The raw route dataset contains 5038 independent trips from November 2019 to March 2020. The study is based on the secondary data from a previous study (Qiu and Chang 2021); after filtering the rides to those exhibited commuting behaviors using travel speed, distance, duration, location, etc. as criteria, 4430 trips are left. Meanwhile, many previous studies have investigated the cycling behavior of bikeshare users and found that a significant portion was made for commuting purposes (Guo and He 2020, 2021; Qiu and Chang 2021).

Map matching: To overcome the inherent GPS route deviation because of urban canyon and map database issues, we need to match GPS points to street segments (Chen et al. 2020a). The short time interval of raw GPS points along trajectories provides good precision in aggregating the GPS points onto the road segments. The average distance between two points is 12.87 m, much shorter than a local street segment or a street block (average 80.5 m) (Qiu and Chang 2021). A map-matching algorithm was adopted to find the match point of the GPS trajectory point on exploded road segments from Open Street Map (OSM) road network shapefile (Chen et al. 2020a). The principle of the algorithm is to draw perpendiculars to road lines and match the point to the foot of the perpendicular with the least length for it is the most possible position on the road grid. We have 32,649 road segmentations with unique IDs ranging from 1 to 32,649.

$$P_t^{'} = argmin(verticaldis(P_t, R_i)), i/in\{1, 2, ..., 32649\}$$

where, P_t is the raw GPS point; $P_t^{'}$ is the matched point on the road grid; R_i is the road segmentation identity; $verticaldis$ is the function that computes the vertical distance between the raw GPS point and road segmentation.

Moreover, we manually validated the accuracy rate of our GPS data remapping framework by randomly selecting 20 GPS routes (see Appendix Fig. 8.5). The results are highly reliable: within 20 random samples, 18 (90%) GPS routes have been perfectly remapped to the correct road network, indicating the effectiveness of the remapping framework.

Trip volume computation: We applied the topology algorithm to all road segmentations. The road network can be represented as a graph structure G. Each road segmentation is an edge of the graph. Each intersection point of the road segmentations is a node of the graph and is given a unique ID according to the logical connection of road segmentations. Then, we use Dijkstra's algorithm to find the shortest path between two matched GPS points. And because the GPS raw dataset has relatively high accuracy, the shortest search function lets us properly define the trip trajectory along the road segment. Dijkstra's algorithm has been widely used in relevant urban studies to find the shortest path between any two vertices of a graph (Schulz et al. 2001; Risald and Suyoto 2017; Salazar Miranda et al. 2021). In our research, adopting Dijkstra's algorithm has very minimal impact on how people choose a specific road with special characteristics. Finally, the trip volume of each unique road segmentation is calculated according to the number of routes traveled on it. The trip volume of each street reflects the desirability for cyclists to travel and is the dependent variable for later regression models.

8.3.3.2 Independent Variables: Micro-level Environment Characteristics

Objective Features

The objective street features were calculated through the following steps.

Downloading Google SVIs: Previous research (Ito and Biljecki 2021) has shown that an SVI sampled from the street centerline could proxy a biker's view and reflects the human perceived environment. Thus, the micro-level objective features extracted from SVIs can be seen as close proxies of perceived reality. In this research, SVIs were obtained through the Google Street View Static API. Following previous research, SVIs were sampled every 50 m along the street segment in QGIS (Song et al. 2022c,) and by feeding the obtained coordinates, each sample point's SVI could be downloaded. The camera settings (heading: parallel with the street, field of view: 120 degrees, the pitch angle: 0 degree) were kept consistent for the study, each SVI is 600×400 pixels, and we obtained 11,426 valid SVIs.

Computing Objective Features: The objective features were calculated using View Index. The view index represents the percentage of the pixels of a specific visual element in an SVI (Zhang et al. 2018). The index of each objective feature in the SVIs was calculated as follows:

$$VI_{fea} = \frac{\sum_{i=1}^{n} PIXEL_{fea}}{\sum_{i=1}^{m} PIXEL_{total}}$$

where VI_{Fea} is the view index of one feature fea in one SVI, $\sum_{i=1}^{m} PIXEL_{total}$ is the total pixels of one SVI, and $\sum_{i=1}^{n} PIXEL_{fea}$ is the pixels of streetscape feature fea in one SVI.

Pyramid Scene Parsing Network (PSPNet) was adopted to calculate the view index representing each type of objective feature. PSPNet has made notable contributions to semantic segmentation (Zhao et al. 2017) and has achieved more than 93.4% pixel-level accuracy. The algorithm has been adopted in numerous relevant urban studies (Verma et al. 2020; Tian et al. 2021; Qiu et al. 2022, 2023; Wang et al. 2022; Xu et al. 2022; Song et al. 2023; He et al. 2023). This study used the algorithm pre-trained on the widely used ADE20K dataset, which collected urban scenes from more than 50 cities and is particularly suitable for the semantic segmentation of cityscapes (Zhou et al. 2019). Finally, 31 types of objective features were extracted in our study using downloaded SVIs, and they included ashcan, tree, bicycle, streetlight, bridge, building, sky, etc. (See Appendix Table 8.4).

Perceived Qualities

Following our prior research, the ML-based framework we developed was used to evaluate the subjectively measured perceived qualities of street scenes (Tian et al. 2021; Qiu et al. 2022).

Scoring the perceived qualities: We randomly selected 300 SVIs and built a crowdsourcing visual survey collecting seven perceived quality preference comparisons from a panel of experts. The perceived qualities included (1) Order, (2) Accessibility, (3) Aesthetic, (4) Ecology, (5) Enclosure, (6) Richness, and (7) Scale. Through the Microsoft TrueSkill algorithm and normalization, the preference results were translated to ranking scores which ranged from 0 to 1.

Training ML models to predict the perceived qualities: The training data was the 300 SVIs from the online visual survey. Their associated seven ranked scores of perceived qualities were used as output labels, while the objective view indices extracted from these SVIs served as input explanatory variables. We split them by 80% and 20% for training and testing purposes. We chose eight ML algorithms, such as Voting Selection and K-Nearest Neighbors, to predict the seven qualities. Those algorithms were classical and matured in supervised learning. Prediction results were compared by the balance performance judged by R-squared (R2), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). On average, Random Forest (RF) was the best-performing model. We then applied the RF algorithm to predict the perceived qualities for all 11,426 SVIs (See Appendix Table 8.4).

8.3.3.3 Other Independent Variables—Macro-level Characteristics

POI Attributes: POI data was obtained from OSM. POI were categorized as (1) Agricultural, (2) Commercial, (3) Community, (4) Industrial, (5) Public Services, (6) Recreation, and (7) Residential. The POI of each street was calculated by the number within the 50 m buffer along the road segment. We also used Shannon Index to measure the mixture of total POI (*Total_Mix*) and the mixture of commercial and residential POI (*Com_Res_mix*) within the buffer area.

Accessibility Attributes: We chose Line Connectivity (*LConn*) and Mean Angular Distance in Radius (*MAD*) as the matrix to evaluate the accessibility attributes of the street network using Space Syntax. *Lconn* is the number of other ending lines to which one line is connected. *MAD* is the mean angular distance within the 800 m (5-min bike ride distance) radius.

Demographic Attributes: Demographic attributes included population density (*PopDen*) and work population density (*WorkDen*). We aggregated the value of each census division zone to the sampling points of the road segments (See Appendix Table 8.4).

8.3.4 Correlation Analysis

First, Pearson correlation analysis was conducted between the trip volume and each variable. Only those variables demonstrating statistically significant correlations with DBS trip volume were selected for further regression analysis; the insignificant variables were removed. Second, we analyzed each group of attributes using the OLS model separately to compare their group-level explanatory power. Third, a baseline model is constructed using control variables (all macro-level characteristics) (Model 1). Then human perceived micro-level characteristics were added to Model 1 to establish Model 2, Model 3, and Model 4 accordingly. The performance of each model and the impacts of human-perceived street environment features were carefully examined and compared. Table 8.1 provides the architecture of each model.

The model can be written as follows:

$$Y_i = \beta_0 + \sum\nolimits_j \beta_j X_{ij} + \varepsilon_i$$

where Y_i is DBS trip volume at observation i; X_{ij} is independent variable j at observation i; β_j is the coefficient of independent variable j. ε_i is an error term at observation i. β_0 is the intercept.

Table 8.1 OLS model architecture

Model	Dependent variable (Y)	Independent variables (X)
Model 1	Bikeshare trip volume	Macro-level Characteristics (POI + Accessibility + Demographic)
Model 2	Bikeshare trip volume	Macro-level Characteristics (POI + Accessibility + Demographic) + Objective Features
Model 3	Bikeshare trip volume	Macro-level Characteristics (POI + Accessibility + Demographic) + Perceived Qualities
Model 4	Bikeshare trip volume	Macro-level Characteristics (POI + Accessibility + Demographic) + Objective Features + Perceived Qualities

8.4 Results

8.4.1 Trip Volumes

Figure 8.2 shows the spatial distribution of the bikeshare trip volume of each road segment. More cycling trip volume could be observed in Cornell Campus, Collegetown, and Downtown.

Figure 8.3a and b show the example of spatial distributions of objective features, streetlight, and tree. In comparison, Fig. 8.4a and b demonstrate the spatial pattern of perceived qualities using Aesthetics and Scale as examples. The heterogeneity regarding the distributions of micro-level environment characteristics can be clearly observed.

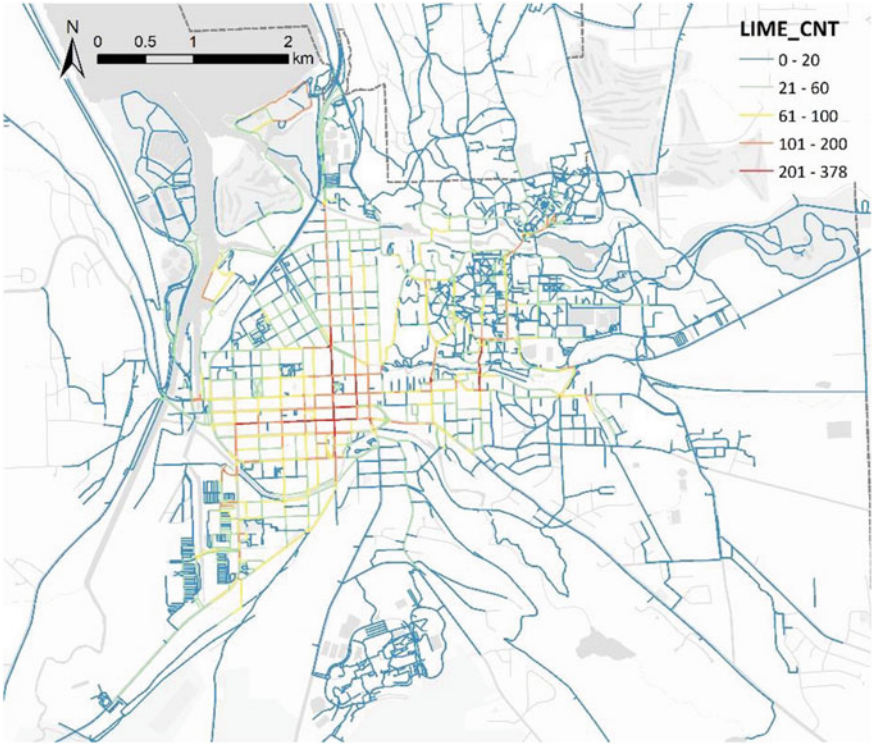


Fig. 8.2 Spatial distribution of the DBS trip volume in the city of Ithaca

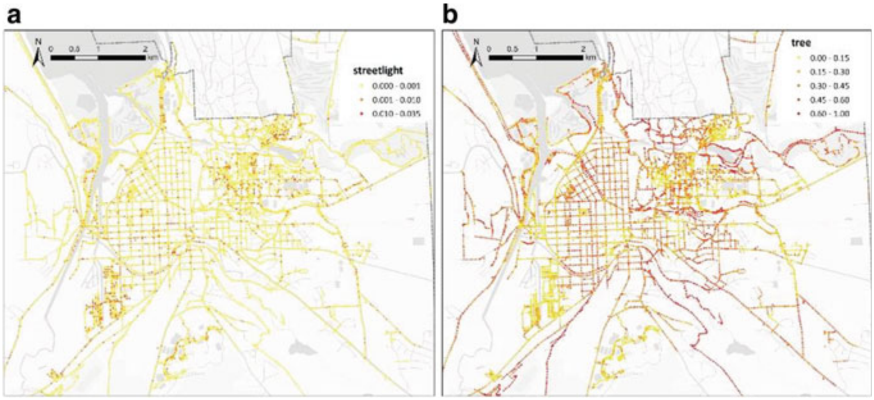


Fig. 8.3 Example of objective features. **a** left: streetlight; **b** right: tree

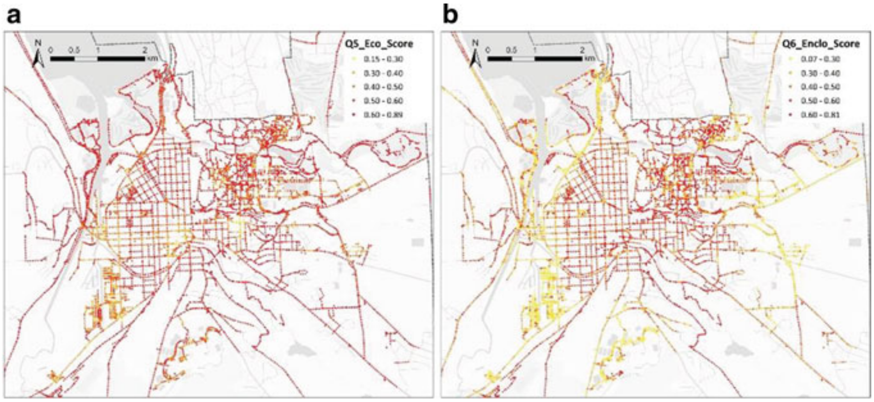


Fig. 8.4 Example of perceived qualities. **a** left: ecology; **b** right: enclosure

8.4.2 Strength of Association by Attribute Groups

We conducted the Pearson correlation analysis between the bikeshare cycling trip volume and all the variables. Only those variables which show statistically significant correlations ($p < 0.05$) with cycling trip volume were kept for further analysis. We then analyzed each group of remaining variables using the OLS model separately. Table 8.2 shows the five attribute groups' explanatory power using R2, and they rank as: POI attributes (0.405) > perceived qualities (0.394) > objective features (0.371) > accessibility attributes (0.266) > demographic attributes (0.172). All five attribute groups pass the F-statistic test ($p < 0.01$), confirming each group's significant role in affecting cycling trip volume. The White Test shows the p-value as 0.28, confirming the absence of heteroscedasticity. The Q-Q probability plot further indicates that the errors were normally distributed.

Table 8.2 Overall regression performance of all OLS model

Model	Model 1	Model 2	Model 3	Model 4
Attributes	Baseline	Baseline + objective features	Baseline + perceived qualities	Baseline + objective features + perceived qualities
Method	OLS	OLS	OLS	OLS
R2	0.448	0.478	0.475	0.488
Adjusted R2	0.447	0.477	0.474	0.487
Prob (F-statistic)	0.000***	0.000***	0.000***	0.000***

Notes p value * < 0.1, ** p < 0.05, *** p < 0.01

It also should be noted that due to the spatial nature of the data presented in this study, the global Moran's I value of 0.55 regarding the OLS residue confirmed the existence of strong spatial autocorrelation. Thus, spatial statistics and Geographically Weighted Regression (GWR) are more appropriate. The authors did fit a spatial model but decided not to present it here, given the main focus, the complexity, and the length of this short chapter. The spatial model results are available upon request.

8.4.3 OLS Regression Results and Performances

We constructed a baseline model (Table 8.2) using macro-level characteristics, including POI attributes, accessibility attributes, and demographic attributes variables. We deleted several variables to avoid the multicollinearity issue. Six POI attribute variables, one accessibility attribute variable, and one demographic attribute variable remained (Model 1). Then different types of micro-level environment characteristics were added to Model 1 to establish Model 2, Model 3, and Model 4 accordingly. It is worth noting that objective features and perceived qualities have been reported to partially overlap in previous research when they explain housing prices (Qiu et al. 2022). We ran the Pearson correlation and calculated Variance Inflation Factor (VIF) between these micro-scale environment variables. Those variables exhibiting low correlation and high multicollinearity (VIF > 10) (Qiu et al. 2023) were removed, and eight objective features and five perceived qualities were retained and used in this research in Model 2, 3, and 4. Nevertheless, it is worth noting that after removing several variables, the retained perceived qualities variables still exhibited higher VIF values, which are caused by the inclusion of powers or products effect of objective feature variables. In summary, our results comply with the VIF criteria, except for the perceived qualities variables. But there were no adverse consequences from multicollinearity in this case (Allison 1977, 2012; Cuadrado-Ballesteros et al. 2017), and it didn't affect the significance, sign of the coefficient and relative magnitude, therefore, the study safely ignored the multicollinearity.

The result shows that: (1) The explanatory power of Model 2, Model 3, and Model 4 were improved compared to Model 1, indicating the effectiveness of micro-level environment characteristics (objective features or perceived qualities) to improve the prediction ability collectively. Taking Model 2 and Model 3 as examples, the incorporation of micro-level characteristics, such as objective features or subjective perceptual scores, yielded a better explanatory power compared to relying solely on traditional macro-level variables. (2) Model 4 demonstrated the cumulative impact of utilizing both measures of micro-level characteristics, resulting in a greater effect than relying on a single measure. This finding suggests that the objective street elements and perceived qualities quantified in our study partially overlap but can complement each other (Qiu et al. 2022). For example, the perceived quality 'enclosure' was removed due to the multicollinearity problem. But enclosure correlated highly with several objective features (sky, building, road).

Table 8.3 shows the correlation coefficients between bikeshare trip volume and the selected variables from various attribute groups. The regression results demonstrate that the micro-level environment characteristics significantly contributed to bike route choices, and they outperformed macro-level characteristics in their predictive power. All the variables from the perceived qualities and objective features (except the objective feature 'tree') exhibit much higher strengths than all variables in the macro-level characteristics.

Table 8.3 Regression performance of all variables and diagnosis for all OLS models

Model	Model 1		Model 2		Model 3		Model 4		
Attributes	Baseline		Baseline + objective features		Baseline + perceived qualities		Baseline + objective features + perceived qualities		
OLS diagnosis	Coef	P > |t|	Coef	P > |t|	Coef	P > |t|	Coef	P > |t|	VIF
POI attributes									
Commercial	4.8336	0.0000	4.1232	0.0000	3.7694	0.0000	3.6946	0.0000	1.72
Community	7.7659	0.0000	6.4200	0.0000	5.4932	0.0000	5.5252	0.0000	1.23
Public services	7.4612	0.0007	6.0773	0.0049	4.9270	0.0224	4.6152	0.0310	1.03
Recreation	10.5403	0.0000	10.6919	0.0000	10.5658	0.0000	10.3598	0.0000	1.05
Residential	0.6831	0.0000	0.5984	0.0000	0.5657	0.0000	0.4455	0.0000	1.64
Com_Res_mix	5.7542	0.0000	5.1890	0.0000	4.8196	0.0001	3.3958	0.0054	1.79
Accessibility Attributes									
Lconn	3.6127	0.0000	1.6027	0.0000	1.0077	0.0003	1.0329	0.0003	7.46
Demographic attributes									
WorkDen	0.0002	0.0000	0.0001	0.0000	0.0002	0.0000	0.0001	0.0000	1.08

(continued)

Table 8.3 (continued)

Model	Model 1		Model 2		Model 3		Model 4		
Attributes	Baseline		Baseline + objective features		Baseline + perceived qualities		Baseline + objective features + perceived qualities		
OLS diagnosis	Coef	P > |t|	Coef	P > |t|	Coef	P > |t|	Coef	P > |t|	VIF
Objective features									
Building	/	/	22.5237	0.0000	/	/	17.3432	0.0065	1.82
Railing	/	/	144.1403	0.0004	/	/	107.7123	0.0079	1.02
Road	/	/	57.5809	0.0000	/	/	44.2920	0.0000	5.46
Sidewalk	/	/	109.5906	0.0000	/	/	59.2922	0.0000	1.45
Signboard	/	/	188.5750	0.0038	/	/	128.0367	0.0536	1.13
Sky	/	/	-28.6037	0.0000	/	/	−18.3540	0.0001	5.30
Streetlight	/	/	765.6655	0.0000	/	/	438.5443	0.0111	1.17
Tree	/	/	7.7673	0.0000	/	/	8.7790	0.0507	3.73
Perceived qualities									
Q1 order	/	/	/	/	31.6048	0.0000	22.8155	0.0012	
Q2 accessibility	/	/	/	/	41.5774	0.0000	29.0449	0.0000	
Q3 aesthetic	/	/	/	/	−41.0859	0.0000	−43.4987	0.0000	
Q4 ecology	/	/	/	/	−24.5528	0.0000	−31.0669	0.0000	
Q7 scale	/	/	/	/	26.4052	0.0000	40.4113	0.0000	

8.5 Discussion

8.5.1 Micro-level Environment Characteristics

The objective features exhibit large variances in strengths, but overall show the highest strengths compared to other groups of variables. The streetlight, signboard, and railing rank the top 3 features that report a positive effect on bike volume, which can be jointly explained by the presence of critical cycling infrastructures such as traffic barriers, which provides necessary protection and clear wayfinding strategies to help cyclists safely navigate through the city. Collectively they can be interpreted as a proxy of perceived safety which was reported to be an essential factor in choosing the route (Van Dyck et al. 2012; Kerr et al. 2016). Furthermore, the fact that road and sidewalk positively correlate with cycling routes aligns with prior studies. Higher pixels of the road in the street scene imply a wider street profile and interface, allowing for dedicated bike lanes (Mertens et al. 2016). Pedestrian-friendly neighborhoods intentionally design the streets to slow down automobile traffic, and we suggest that the sidewalk implies a sustainable travel mode and encourages cycling activities through the area. In addition, consistent with prior studies, trees contributed to cycling behavior positively, which can be explained by their value of microclimate control, shade provision, and aesthetics (Lu et al. 2019; Porter et al. 2020). And not surprisingly, the building and sky revealed a contrasting effect. The sky is the only element that reported a negative relationship with trip volume in the objective features group. In the North American context, a higher building ratio and less sky ratio imply compact development projects which provide more convenient urban amenities (Fraser and Lock 2011).

Within the perceived qualities group, scale, accessibility, and order positively contribute to the cycling volume. Scale refers to the concept of human scale in that the street features in the physical environment match the proportions of humans, and the result indicates the importance of human scale design in the built environment. Perceived accessibility refers to the people's reading of the landscape so that they can reach geographically dispersed destinations, which implies good visibility of the road layout design. Specifically, scholars have found that cyclists' perceptions of destinations and access can increase the site's attractiveness, in our case, they were likely to increase cycling trip volume (Guo and He 2021). Nevertheless, perceived qualities, including aesthetics and ecology, were negatively correlated with cycling trip volume, which contradicts previous research and our assumptions (Lu et al. 2019; Wang et al. 2020). The possible interpretation is likely due to safety concerns related to overgrown planting, which can block the sight triangles in the road intersections and may become a nuisance to cyclists. Since the results contrast with objective features, better definitions of the concepts should be developed in the future to understand the possible reasons for this discrepancy.

8.5.2 Conventional Macro-level Environment Characteristics

The study further strengthens our knowledge of macro-level characteristics and their associations with cycling behaviors. Internally among all the POI attributes, recreational service is reported to have the highest strength, followed by community service and commercial land use, while residential land use exhibits the lowest strength. This finding is in line with previous built environment research that states the importance of non-residential land use in influencing cycling behavior (Lu et al. 2019; Porter et al. 2020). Lu et al. (2019) has revealed that the objective greenery extracted from eye level using SVI data showed higher strength than the number of retail shops affecting aggregated cycling behavior, which is further supported by our findings at the street segment level. Furthermore, the commercial and residential land use mix (*Com_Res_mix*) was positively correlated with cycling trip volume, which is consistent with prior research (Kerr et al. 2016). Regarding connectivity, it contributed positively to cycling trip volume, although its statistical strength was relatively lower compared to most other variables. Moreover, the sociodemographic attributes showed the most insufficient power among all variables, which resonates with previous research where no significant relationship has been established (Mateo-Babiano et al. 2016; Lu et al. 2019).

8.6 Conclusion

In summary, using an unprecedented amount of GPS bike trajectory data collected from the DBS system in Ithaca, this study was able to aggregate cycling trip volume at the street segment level, and explain its correlation with street desirability. Although a few researches have examined the impact of micro-level environment on cycling behavior through surveys or emerging SVIs, they tend to be constrained to a single objective feature, with rare mentions of the overall perceived qualities. By contrast, this paper provides a useful example of surfacing the impact of perceived qualities on cycling trip volume. Using several OLS regression models, the strengths of both objective features and perceived qualities measured from SVI data were comprehensively analyzed, with their impacts on informing cycling trip volume at the street level explained. First, it was found that the overall regression performance of the OLS model was improved after adding micro-level characteristics to the Baseline model, which only contained macro-level characteristics, and the overall strength of objective features was comparable to perceived qualities. Second, when examining the predictive power of individual variables, both types of micro-level environment characteristics significantly affected the cycling trip volume more than the variables of conventional macro-level characteristics (e.g., POIs). Third, objective features jointly and individually exhibited higher strengths than perceived qualities; for example, streetlight, signboard, and railing exhibited the highest strengths among all the variables. However, we argue that perceived qualities have the value to complement objective features because the model that incorporated both measures exhibited the highest statistical strength. The research proved the significant impacts of micro-level environment characteristics on cycling route choices, which validated the effectiveness of our methodology. It provides a valuable reference to policymakers, urban planners, and operating companies to work together towards a sustainable, cycling-friendly, and vibrant city through the appropriate planning and management of the perceived built environment. For example, to encourage active travel, especially cycling, streetscape designers and engineers should pay attention to the proper placement of streetlights.

This study contained several limitations that can be improved in future studies. First, this study didn't capture how the temporal variance of trip volume is influenced by the built environment variables at the street level (Guo and He 2020). It is worth further investigating how time and seasonality affect route choices. Second, we acknowledge that although we saw improvements in the performances of OLS models after adding micro-level characteristics as variables, the improvements were minor. We plan to disclose those unexplained built environment variables affecting cycling trip volume. Third, cycling behavior is often associated with specific purposes. For example, the bikeshare system has been used as a feeder mode for metro or bus transit (Guo and He 2021; Qiu and Chang 2021). Future research can investigate how perceived characteristics affect trips, particularly for leisure or work purposes (Porter et al. 2020). Fourth, the perceived qualities measured in this study were not comprehensive enough; for instance, they didn't include psychological qualities such as perceived safety which may have a high impact. And a more complex statistical process can be applied to the variables, such as using Principal Component Analysis to reduce perceptual dimensions and explain their combined effects (Song et al. 2022b). Fifth, the training set can be further expanded. For example, more pairwise selections can be made on increased SVIs, and the accuracy can be improved. Sixth, we did not have access to a baseline level of cycling trip volume. Therefore, we could not determine whether the significant characteristics (e.g., streetlight) merely attracted a fixed number of cycling trips to those road segments, or if they actually generated more cycling trips. Seventh, it should be noted that given the spatial nature of the data, we plan to interpret the relationship using appropriate spatial statistics techniques such as GWR. Eighth, we acknowledge that the multicollinearity issue between perceived qualities and objective features should be examined in detail and resolved in the future. Lastly, PSPNet has limitations as it predicts the features based on colors and pixels. The accuracy can be further enhanced by improving the algorithm (Yin et al. 2022) or expanding the cityscape training dataset.

Acknowledgements The authors would like to thank Alan Black Transportation Related Grant (2020, 2021) from the Department of City and Regional Planning and the Serve in Place Fund (2020, 2021) from the Office of Engagement Initiatives at Cornell University for their support. The authors would also like to thank the City of Ithaca, Hector Chang, Bike Walk Tompkins, and Lime for providing secondary bikeshare trip data. This research could not have been made without the support from Bike Walk Tompkins.

Appendix

See Table 8.4 and Fig. 8.5.

Table 8.4 Descriptive statistics of variables

Attributes	Count	Mean	Std	Min	Max
Dependent variable					
Lime bikeshare trip volume	11,416	25.5801	40.27762	1	378
Macro-level characteristics					
POI attributes					
Agricultural	986	–	–	–	–
Commercial	2429	–	–	–	–

(continued)

Table 8.4 (continued)

Attributes	Count	Mean	Std	Min	Max
Community	695	–	–	–	–
Industrial	94	–	–	–	–
Public services	178	–	–	–	–
Recreation	181	–	–	–	–
Residential	24,390	–	–	–	–
Total_Mix	11,426	0.2437	0.40691	0	1
Com_Res_mix	11,426	0.1513	0.30965	0	1
Accessibility attributes					
Lconn	32,649	2.6076	0.9610	1	10
Lnk800c	32,649	1265.6442	476.0376	750.3320	5758.8901
Demographic attributes					
PopDensity	11,426	2173.286	3885.097	0	27,896.44
WorkDensity	11,426	1894.722	12,028.57	0	176,883.4
Micro-level characteristics					
Objective features					
Ashcan	11,426	0.0000	0.0006	0	0.0250
Awning	11,426	0.0001	0.0011	0	0.0415
Bicycle	11,426	0.0001	0.0012	0	0.0599
Booth	11,426	0.0000	0.0000	0	0.0006
Bridge	11,426	0.0006	0.0083	0	0.3080
Building	11,426	0.0485	0.0883	0	0.9708
Ceiling	11,426	0.0004	0.0121	0	0.4811
Chair	11,426	0.0000	0.0003	0	0.0223
Column	11,426	0.0001	0.0025	0	0.1788
Earth	11,426	0.0231	0.0610	0	0.8837
Fence	11,426	0.0050	0.0178	0	0.3732
Fountain	11,426	0.0000	0.0002	0	0.0182

(continued)

Table 8.4 (continued)

Attributes	Count	Mean	Std	Min	Max
Grass	11,426	0.1090	0.0950	0	0.7059
Lake	11,426	0.0000	0.0004	0	0.0409
Lamp	11,426	0.0000	0.0000	0	0.0001
Minibike	11,426	0.0000	0.0008	0	0.0407
Mountain	11,426	0.0015	0.0224	0	0.9219
Person	11,426	0.0004	0.0028	0	0.1228
Plant	11,426	0.0144	0.0364	0	0.5689
Railing	11,426	0.0007	0.0078	0	0.2369
Road	11,426	0.2136	0.1155	0	0.5163
Sculpture	11,426	0.0000	0.0004	0	0.0326
Sidewalk	11,426	0.0182	0.0406	0	0.4463
Signboard	11,426	0.0017	0.0049	0	0.1299
Sky	11,426	0.1890	0.1242	0	0.4881
Streetlight	11,426	0.0006	0.0019	0	0.0352
Car	11,426	0.0152	0.0296	0	0.2698
Tree	11,426	0.3192	0.1688	0	0.9967
Van	11,426	0.0003	0.0044	0	0.2072
Wall	11,426	0.0030	0.0155	0	0.6201
Water	11,426	0.0010	0.0095	0	0.3042
Perceived qualities					
Q1_Order	11,426	0.4864	0.0564	0.2798	0.6988
Q2_Access	11,426	0.4694	0.1090	0.2008	0.8279
Q3_Beauty	11,426	0.5331	0.0921	0.2653	0.7257
Q4_Ecology	11,426	0.5949	0.1271	0.1452	0.8929
Q5 Enclosure	11,426	0.4662	0.1744	0.0671	0.8055
Q6_Richness	11,426	0.4279	0.1018	0.2338	0.7052
Q7_Scale	11,426	0.4732	0.0923	0.2261	0.7272

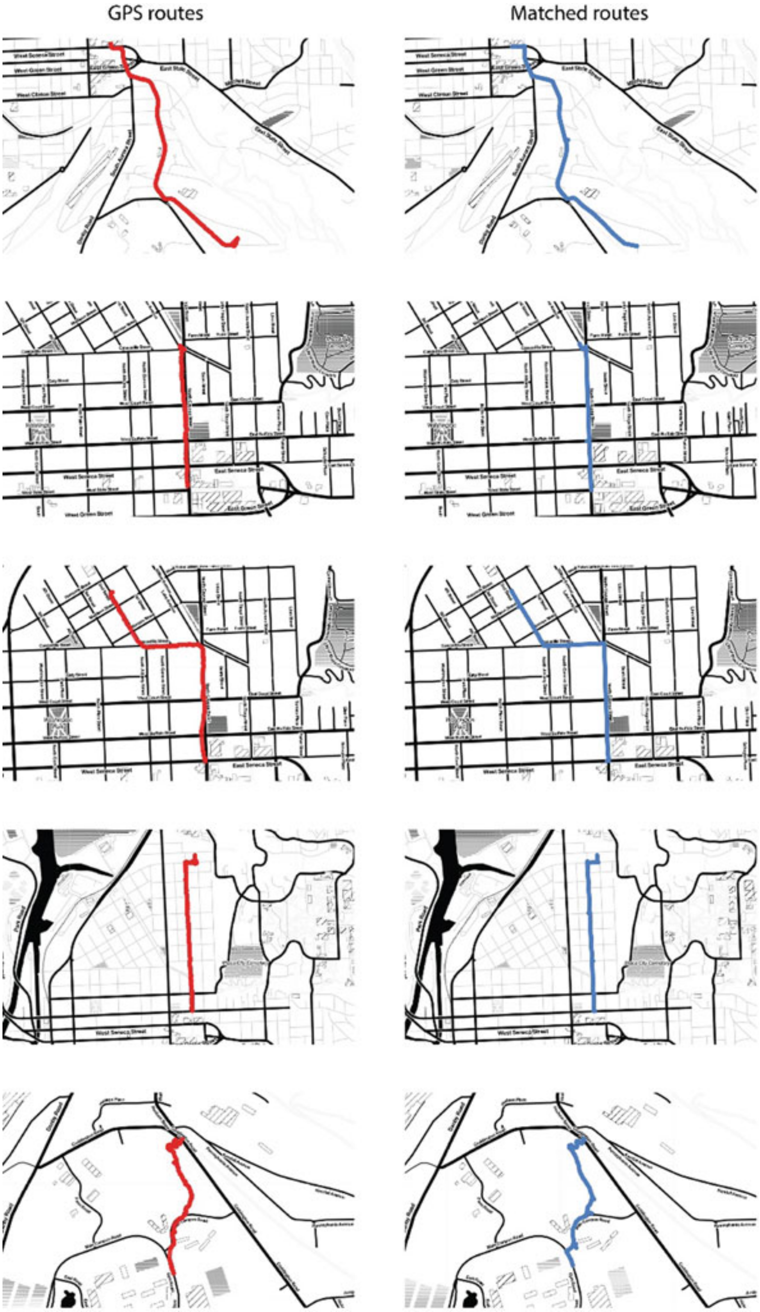


Fig. 8.5 Examples of GPS from the randomly selected samples and matched routes for validation

References

Allison P (2012) When can you safely ignore multicollinearity. Stat Horiz 5:1–2

Allison PD (1977) Testing for interaction in multiple regression. Am J Sociol 83:144–153. https://doi.org/10.1086/226510

Aultman-Hall L, Hall FL, Baetz BB (1997) Analysis of bicycle commuter routes using geographic information systems: implications for bicycle planning. Transp Res Rec 1578:102–110. https://doi.org/10.3141/1578-13

Chen J, Li W, Zhang H et al (2020a) Mining urban sustainable performance: GPS data-based spatio-temporal analysis on on-road braking emission. J Clean Prod 270:122489

Chen Z, van Lierop D, Ettema D (2020b) Dockless bike-sharing systems: what are the implications? Transp Rev 40:333–353. https://doi.org/10.1080/01441647.2019.1710306

Cuadrado-Ballesteros B, Martínez-Ferrero J, García-Sánchez IM (2017) Mitigating information asymmetry through sustainability assurance: the role of accountants and levels of assurance. Int Bus Rev 26:1141–1156. https://doi.org/10.1016/j.ibusrev.2017.04.009

Dubey A, Naik N, Parikh D et al (2016) Deep learning the city: quantifying urban perception at a global scale. In: Leibe B, Matas J, Sebe N, Welling M (eds) Computer vision—ECCV 2016. Springer International Publishing, Cham, pp 196–212

Ewing R, Handy S (2009) Measuring the unmeasurable: urban design qualities related to walkability. J Urban Des 14:65–84. https://doi.org/10.1080/13574800802451155

Fraser SDS, Lock K (2011) Cycling for transport and public health: a systematic review of the effect of the environment on cycling. Eur J Pub Health 21:738–743. https://doi.org/10.1093/eurpub/ckq145

Gong F-Y, Zeng Z-C, Zhang F et al (2018) Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. Build Environ 134:155–167. https://doi.org/10.1016/j.buildenv.2018.02.042

Goodman A, Sahlqvist S, Ogilvie D (2014) New walking and cycling routes and increased physical activity: one- and 2-year findings from the UK iConnect study. Am J Public Health 104:e38–e46. https://doi.org/10.2105/AJPH.2014.302059

Goodspeed R, Yan X (2017) Crowdsourcing street beauty: visual preference surveys in the big data era. In: Schintler L, Chen Z (eds) Big data for regional science, 1st edn. Routledge, London and New York, pp 75–93

Gu T, Kim I, Currie G (2019) To be or not to be dockless: empirical analysis of dockless bikeshare development in China. Transp Res Part A: Policy Pract 119:122–147. https://doi.org/10.1016/j.tra.2018.11.007

Guo Y, He SY (2020) Built environment effects on the integration of dockless bike-sharing and the metro. Transp Res Part D: Transp Environ 83:102335. https://doi.org/10.1016/j.trd.2020.102335

Guo Y, He SY (2021) Perceived built environment and dockless bikeshare as a feeder mode of metro. Transp Res Part D: Transp Environ 92:102693. https://doi.org/10.1016/j.trd.2020.102693

He J, Zhang J, Yao Y, Li X (2023) Extracting human perceptions from street view images for better assessing urban renewal potential. Cities 134:104189. https://doi.org/10.1016/j.cities.2023.104189

Ito K, Biljecki F (2021) Assessing bikeability with street view imagery and computer vision. Transp Res Part C: Emerg Technol 132:103371. https://doi.org/10.1016/j.trc.2021.103371

Kang Y, Zhang F, Gao S et al (2021) Human settlement value assessment from a place perspective: considering human dynamics and perceptions in house price modeling. Cities 118:103333. https://doi.org/10.1016/j.cities.2021.103333

Kerr J, Emond JA, Badland H et al (2016) Perceived neighborhood environmental attributes associated with walking and cycling for transport among adult residents of 17 cities in 12 countries: the IPEN study. Environ Health Perspect 124:290–298. https://doi.org/10.1289/ehp.1409466

Krenn PJ, Oja P, Titze S (2014) Route choices of transport bicyclists: a comparison of actually used and shortest routes. Int J Behav Nutr Phys Act 11:31. https://doi.org/10.1186/1479-5868-11-31

Law S, Sakr FL, Martinez M (2014) Measuring the changes in aggregate cycling patterns between 2003 and 2012 from a space syntax perspective. Behav Sci 4:278–300. https://doi.org/10.3390/bs4030278

Li X, Zhang C, Li W et al (2015) Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For Urban Green 14:675–685. https://doi.org/10.1016/j.ufug.2015.06.006

Lin L, Moudon AV (2010) Objective versus subjective measures of the built environment, which are most effective in capturing associations with walking? Health Place 16:339–348. https://doi.org/10.1016/j.healthplace.2009.11.002

Lu Y, Yang Y, Sun G, Gou Z (2019) Associations between overhead-view and eye-level urban greenness and cycling behaviors. Cities 88:10–18. https://doi.org/10.1016/j.cities.2019.01.003

Ma L, Cao J (2019) How perceptions mediate the effects of the built environment on travel behavior? Transportation 46:175–197. https://doi.org/10.1007/s11116-017-9800-4

Ma L, Dill J (2017) Do people's perceptions of neighborhood bikeability match "reality"? J Transp Land Use 10:291–308

Mateo-Babiano I, Bean R, Corcoran J, Pojani D (2016) How does our natural and built environment affect the use of bicycle sharing? Transp Res Part A: Policy Pract 94:295–307. https://doi.org/10.1016/j.tra.2016.09.015

Mertens L, Compernolle S, Gheysen F et al (2016) Perceived environmental correlates of cycling for transport among adults in five regions of Europe. Obes Rev 17:53–61. https://doi.org/10.1111/obr.12379

Porter AK, Kohl HW, Pérez A et al (2020) Bikeability: assessing the objectively measured environment in relation to recreation and transportation bicycling. Environ Behav 52:861–894. https://doi.org/10.1177/0013916518825289

Qiu W, Chang H (2021) The interplay between dockless bikeshare and bus for small-size cities in the US: a case study of Ithaca. J Transp Geogr 96:103175. https://doi.org/10.1016/j.jtrangeo.2021.103175

Qiu W, Li W, Liu X et al (2023) Subjective and objective measures of streetscape perceptions: relationships with property value in Shanghai. Cities 132:104037. https://doi.org/10.1016/j.cities.2022.104037

Qiu W, Zhang Z, Liu X et al (2022) Subjective or objective measures of street environment, which are more effective in explaining housing prices? Landsc Urban Plan 221:104358

Raford N, Chiaradia A, Gil J (2007) Space syntax: the role of urban form in cyclist route choice in Central London

Risald, Mirino AE, Suyoto (2017) Best routes selection using Dijkstra and Floyd-Warshall algorithm. In: 2017 11th international conference on information and communication technology and system (ICTS). pp 155–158

Salazar Miranda A, Fan Z, Duarte F, Ratti C (2021) Desirable streets: using deviations in pedestrian trajectories to measure the value of the built environment. Comput Environ Urban Syst 86:101563. https://doi.org/10.1016/j.compenvurbsys.2020.101563

Schulz F, Wagner D, Weihe K (2001) Dijkstra's algorithm on-line: an empirical case study from public railroad transport. ACM J Exp Algorithmics 5:12-es. https://doi.org/10.1145/351827.384254

Song Q, Dou Z, Qiu W et al (2023) The evaluation of urban spatial quality and utility trade-offs for Post-COVID working preferences: a case study of Hong Kong. ARIN 2:1. https://doi.org/10.1007/s44223-022-00020-x

Song Q, Li M, Qiu W, et al (2022a) The coherence and divergence between the objective and subjective measurement of street perceptions for Shanghai. In: Chen W, Yao L, Cai T, et al (eds) Advanced data mining and applications. Springer Nature Switzerland, Cham, pp 244–256

Song Q, Li W, Li M, Qiu W (2022b) Social Inequalities in neighborhood-level streetscape perceptions in Shanghai: the coherence and divergence between the objective and subjective measurements

Song Q, Liu Y, Qiu W et al (2022c) Investigating the impact of perceived micro-level neighborhood characteristics on housing prices in Shanghai. Land 11:2002. https://doi.org/10.3390/land11112002

Su N, Li W, Qiu W (2023) Measuring the associations between eye-level urban design quality and on-street crime density around New York subway entrances. Habitat Int 131:102728. https://doi.org/10.1016/j.habitatint.2022.102728

Tarpin-Pitre L, Morency C (2020) Typology of bikeshare users combining bikeshare and transit. Transp Res Rec 2674:475–483. https://doi.org/10.1177/0361198120936262

Tian H, Han Z, Xu W, et al (2021) Evolution of historical urban landscape with computer vision and machine learning: a case study of Berlin. JoDLA 436–451

Van Dyck D, Cerin E, Conway TL et al (2012) Perceived neighborhood environmental attributes associated with adults' transport-related walking and cycling: findings from the USA, Australia and Belgium. Int J Behav Nutr Phys Act 9:70. https://doi.org/10.1186/1479-5868-9-70

Verma D, Jana A, Ramamritham K (2020) Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. Build Environ 186:107340. https://doi.org/10.1016/j.buildenv.2020.107340

Wang R, Lu Y, Wu X et al (2020) Relationship between eye-level greenness and cycling frequency around metro stations in Shenzhen, China: a big data approach. Sustain Cities Soc 59:102201. https://doi.org/10.1016/j.scs.2020.102201

Wang X, Liu Y, Zhu C et al (2022) Associations between the streetscape built environment and walking to school among primary schoolchildren in Beijing, China. J Transp Geogr 99:103303. https://doi.org/10.1016/j.jtrangeo.2022.103303

Wang Y, Qiu W, Jiang Q et al (2023) Drivers or pedestrians, whose dynamic perceptions are more effective to explain street vitality? A case study in Guangzhou. Remote Sensing 15:568. https://doi.org/10.3390/rs15030568

Winters M, Teschke K, Grant M et al (2010) How far out of the way will we travel?: Built environment influences on route selection for bicycle and car travel. Transp Res Rec 2190:1–10. https://doi.org/10.3141/2190-01

Xu X, Qiu W, Li W, et al (2022) Associations between street-view perceptions and housing prices: subjective vs objective measures using computer vision and machine learning techniques. Remote Sensing 14:891. https://doi.org/10.3390/rs14040891

Yin Y, Guo Y, Deng L, Chai B (2022) Improved PSPNet-based water shoreline detection in complex inland river scenarios. Complex Intell Syst. https://doi.org/10.1007/s40747-022-00793-8

Zhang F, Zhou B, Liu L et al (2018) Measuring human perceptions of a large-scale urban region using machine learning. Landsc Urban Plan 180:148–160

Zhao H, Shi J, Qi X, et al (2017) Pyramid scene parsing network. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). pp 6230–6239

Zhou B, Zhao H, Puig X et al (2019) Semantic understanding of scenes through the ADE20K dataset. Int J Comput vis 127:302–321. https://doi.org/10.1007/s11263-018-1140-0

Chapter 9
A Planning Support System for Boosting Bikeability in Seoul

Madiha Bencekri, Donggyun Ku, Doyun Lee, and Seungjae Lee

Abstract The popularity of cycling has increased because it is conceived to be a much safer and environmentally friendly travel mode than automobiles and mass transit. As a result, shared-bike programs have gained attention among academics and policymakers. While inefficacies and other deficits in shared-bike programs have been described in the literature, so have these programs' significant environmental, health, and economic benefits. The authors of this study developed a planning support system (PSS) to boost the efficacy and performance of the shared-cycling system in Seoul, Korea, using a developed bikeability index. First, they employed the Massachusetts Institute of Technology urban network analysis (UNA) and other spatial and statistical analyses to generate the model's input dataset. Second, the authors built the bikeability index using geographically weighted regression (GWR) analysis. Using the index, the authors evaluated Seoul's shared-bike system performance and identified important global and local variables affecting its efficacy. Based on these variables, the authors developed change scenarios. Finally, the authors used the planning support system to simulate these scenarios in multiple iterations until they reached a metric solution using the PSS feedback loop process.

Keywords Shared bike program · Bikeability index · Geographically weighted regression · Feedback loop · Planning support system

M. Bencekri · D. Ku
Department of Transportation Engineering/Department of Smart Cities, University of Seoul, Seoul, Korea
e-mail: koo2646@naver.com

D. Lee
Graduate School of Environmental Studies, Seoul National University, Seoul, Korea
e-mail: un2ver5e@naver.com

S. Lee (✉)
Department of Transportation Engineering, University of Seoul, Seoul, Korea
e-mail: sjlee@uos.ac.kr

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_9

9.1 Introduction

With the rise in environmental awareness, both policymakers and researchers are increasingly interested in the sustainable mobility planning approach, including the promotion of non-motorized and eco-friendly mobility (Mrkajić and Anguelovski 2016). Cycling is considered a sustainable and eco-friendly transportation mode that enhances the first-and-last-mile connection and thereby also improves accessibility and equity (Zuo et al. 2020). Thus, the popularity of shared-bike programs has increased tremendously among the academic and policymaking communities (Eren and Uz 2020). Worldwide, more than 2,000 shared-bike programs were implemented by 2021 (Xin et al. 2022). Often, shared-bike programs are divided into dock-based and dock-less types. The most widely used type is the dock-based program, which consists of picking up the bike from the closest station to the origin and returning it to the closest station to the destination (Xin et al. 2022).

In Seoul, the biking rental system was started in April 2000 with the installation of rental docks near two subway stations, Changdong Station and Yeouido Station, followed by the implementation of a free shared-bicycle service in the Songpa-gu district in 2004 (Fig. 9.1). But officially, Seoul's shared-bike program began in September 2015, when 150 docking stations and 2,000 bicycles were added that can be used easily through a mobile application. The objectives were to mitigate traffic congestion, air pollution, and high oil prices and to improve the quality of life. Since July 2016, the shared-bike program was expanded to reach 2,600 docking stations and 40,500 bicycles by 2021. To support Seoul's shared-bike program, in 2012 the Seoul Metropolitan Government (SMG) launched the "Cycle Rapid Transportation" program, which aimed to create 1,330 km of bike-only traffic lanes by 2030, of which 940 km was completed by 2020 (Fig. 9.1).

Worldwide, shared-bike programs suffer from a lack of efficacy as a result of many challenges, such as transforming cities into bicycle-friendly cultures (Zayed 2016). Therefore, to better understand some of these challenges and boost Seoul's bike-friendliness, we designed a bike planning support system (PSS).

The research employed a comprehensive dataset including (1) the built and natural environment, (2) public transit and bike demand and infrastructure, (3) socioeconomic data, and (4) economic activity. The safety aspect was not considered in this study because of the absence of data and the low rate of cycling accidents, which represent only 6% of road fatalities in Seoul (OECD, 2021). However, a safety variable could be implemented as an additional component of the bike PSS, in further development.

The bike PSS consists of four blocks: (1) creating the dataset and identifying the most important variables, (2) building a bikeability index for the city, (3) assessing the shared-cycling system using the bikeability index and drafting personalized interventions and scenarios, and (4) implementing these various scenarios in the PSS in multiple iterations until reaching the desired bikeability index for the city. The PSS is based on a positive feedback loop of the selected scenarios.

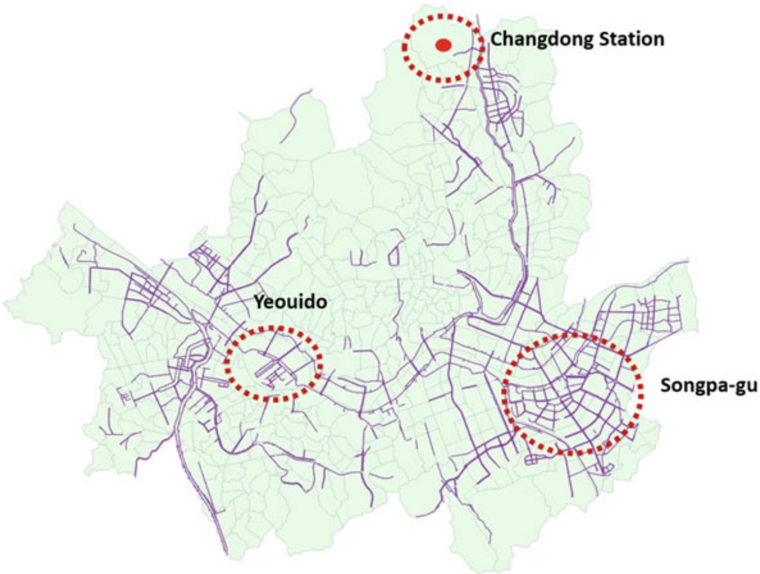


Fig. 9.1 Cycle rapid transportation in Seoul

9.2 Literature Review

9.2.1 Bikeability Index

The term "bikeability" represents the convenience of reaching by bicycle significant locations based on the distance traveled, the perceived comfort, or the attractiveness of the streets (Grigore et al. 2019). In contrast, "bike-friendliness" is a measure of a geographic area's bikeability, safety, and laws that promote biking within a community (Lowry et al. 2012). The bikeability index has emerged as a valuable parameter for improving bikeability and bike-friendliness (Arellana et al. 2020; Kellstedt et al. 2021). The bikeability index has five characteristics that influence cycling in the built environment: bicycle facility availability, bicycle facility quality, street connections, terrain, and land use (Winters et al. 2013). Thus, a deep understanding of these factors is thought to lead to more cycling-friendly policies (Lu et al. 2019), which in turn could facilitate more bikeability and biking-friendly cities.

Several studies have built a bikeability map considering only the built environment variables, like cycling infrastructure, topography, and land use (Hardinghaus et al. 2021; Zhang et al. 2017). Schmid-Querg et al. (2021) investigated and compared the existing bikeability indices to develop an index specific to the city of Munich. The study used bike paths, speed limits, bike stations, and bike-transit integration (Schmid-Querg et al. 2021). The Copenhagen Index revealed five leading indicators of a city's readiness to adopt cycling, which are population, road network length, city form, city area, and modal split of motorized transportation (Zayed 2016).

Altogether, while some researchers have found that the built environment is the key parameter affecting cycling (Shaer et al. 2021), others consider the built environment alone to be an unreliable factor (Hardinghaus et al. 2021). Such studies have found that bikeability is influenced by the built environment of the city (Lee and Bencekri 2021; Shaer et al. 2021), road infrastructure (Yang et al. 2019), socioeconomic characteristics (Mrkajić and Anguelovski 2016; Vidal Tortosa et al. 2021), land-use mix (Jiao et al. 2022; Kamel et al. 2020), public transit infrastructure (Schmid-Querg et al. 2021; Zayed 2016), distance to transit (Cervero and Kockelman 1997), cycling infrastructure (Arellana et al. 2020), accessibility (Rubulotta et al. 2013; Sarlas et al. 2020; Sevtsuk and Mekonnen 2012), and finally typology (Grigore et al. 2019; Porter et al. 2020). Often, communities with high population density, sophisticated land-use mix, open spaces, and retail areas have the highest rates of active travel, particularly cycling (Hankey et al. 2017). Thus, our research considered the following factors: (1) built and natural environment (e.g., a high inclines prevent convenient cycling and willingness to cycle); (2) public transit and cycling demand, and infrastructure, given that cycling is intended to boost the first–last mile connection; (3) population and employment, to investigate the existing potential demand of cycling; (4) the Massachusetts Institute of Technology (MIT) urban network analysis (UNA) centrality indices to reflect the accessibility of bike stations; and (5) economic activity, to investigate its relationship to cycling demand given that retail consumption could indicate the livability and movement within a district.

9.2.2 Planning Support Systems

A planning support system (PSS) is a geo-information-based tool that helps planners with tasks such as communication, analysis, and planning (Vonk et al. 2007). Often, a PSS uses big data, data mining, analysis, visualization, and modeling to ease planning and enhance sustainability (Pettit et al. 2018). The framework of the planning support system combines a methodology to guide analysis, prescription, and prediction; identification of the planning issues; and application of data to inform design and modeling (Geertman and Stillwell 2003).

A study in Amsterdam showed that a PSS is important for strategic land use and transport planning and suggested involving transport planners in earlier phases of the planning process (te Brömmelstroet 2009). In Austria, (Fikar et al. 2018)suggested a PSS that incorporated an agent-based simulation and dynamic solution procedures to route conventional requests. This PSS enabled the investigation of various problem settings and contexts with the goals of facilitating micro-hub operation and maintaining enough cargo bikes. However, incorporating various behavioral factors of drivers needs further investigation (Fikar et al. 2018). Caggiani and Ottomanelli (2012) adopted an optimization-based decision support system with a focus on the redistribution of shared bikes within a bike-sharing system (Caggiani and Ottomanelli

2012). Their aim was to minimize the redistribution costs for bike-sharing operations while achieving higher user satisfaction. However, additional simulations are needed for real-time estimation of multiple vehicle-use parameters (Caggiani and Ottomanelli 2012). No studies have investigated a PSS for boosting bikeability or for enhancing cycling friendliness in a city. Our study fills this gap by introducing a PSS to improve the bikeability of each district of Seoul, using a bikeability index that was developed based on several factors influencing cycling-system performance.

9.2.3 Research Gap and Contribution

A few studies have investigated bikeability using the built environment, public transport, and socioeconomic data. However, these studies have been mainly based on an exhaustive analysis of the literature, or they focused on just a few specific factors (Eren and Uz 2020). Very few researchers have investigated a cycling PSS, and none has focused on enhancing bikeability. Therefore, our study bridged that gap by being the first to develop a district-based bikeability index in the city of Seoul using a comprehensive dataset that includes: (1) socioeconomic factors (population and employment densities), (2) built and natural environment factors (inclines and land-use mix indices), (3) public transit and bike demand and infrastructure (station locations and number, and the average distance to bike stations), (4) economic activity (retail consumption), and (5) the UNA centrality indices to reflect accessibility.

More specifically, our research was the first to (1) include an economic activity through retail consumption factor, (2) include bike-stop centrality indices using the MIT Urban Network Analysis (UNA) tool, (3) consider the Transit Oriented Development (TOD) concept through the number of transit stations and the bike–transit distance, and (4) create a bike PSS to enhance the bikeability index. This PSS could assess the current shared-bike system by assessing bikeability in a district, identifying the main factors that make a district more bike-friendly, and predicting the impact of cycling policies.

9.3 Methodology

9.3.1 Methodology Summary

To create the bikeability index, we adapted the flowchart shown in Fig. 9.2 to investigate the bikeability of Seoul according to each district's administrative boundaries using large data on socioeconomics (Hankey et al. 2017; Vidal Tortosa et al. 2021), the built environment (Kamel et al. 2020; Porter et al. 2020; Sarlas et al. 2020; Shaer et al. 2021), and bike and public transit infrastructure (Arellana et al. 2020; Lee and Bencekri 2021; Schmid-Querg et al. 2021; Zayed 2016).

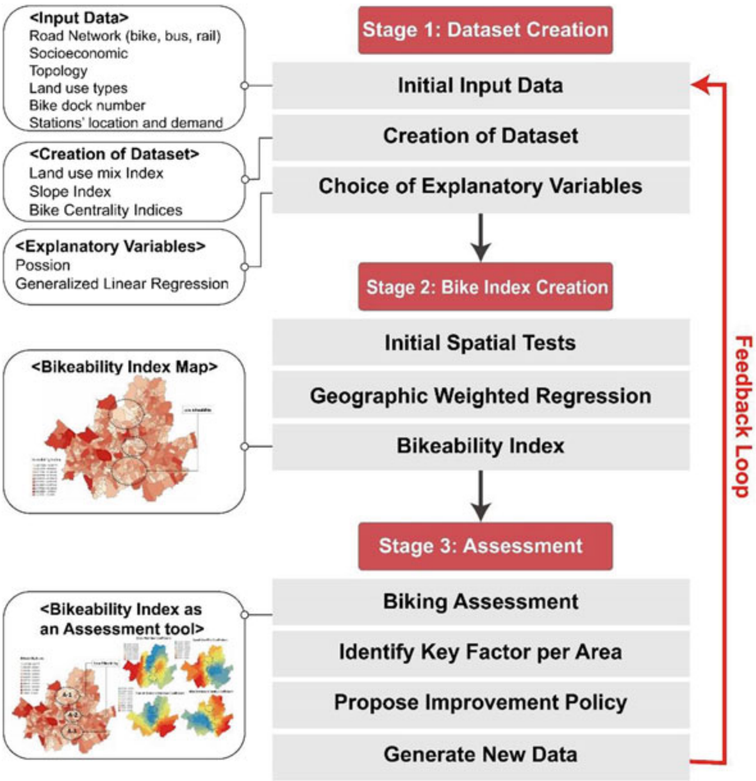


Fig. 9.2 Bike planning support system flow chart

The bikeability index methodology (Fig. 9.2) had four stages. First, we processed data to create a suitable input dataset. Data processing included the computation of (1) the five centrality indices of bike stations, (2) the average shortest distance between bike public transit stations, and (3) the land-use mix index using the entropy theory (Zagorskas 2016). Second, we removed variables with high multicollinearity using Poisson regression because cycling data are of count type and non-Gaussian (Munira et al. 2021). Third, and to justify the use of the geographically weighted regression (GWR), we conducted spatial analyses to prove the spatial autocorrelation, dependence, and heterogeneity using the Moran index, Lagrange multiplier, and Chow test, respectively (Getis 2007; Li et al. 2007). Finally, we developed a GWR model. The GWR is an effective method to explore spatial non-stationarity because it permits a local variation of parameters. Previously, GWR was implemented to investigate bicycle-friendliness, the relationship between cycling and the built environment, cycling behavior, and the impact of cycling on health and the environment (Shin et al. 2022; Wei et al. 2021).

Makarova et al. (2019) suggested a decision support system to plan the development of bicycle infrastructure and to evaluate its safety, proposing possible cycling

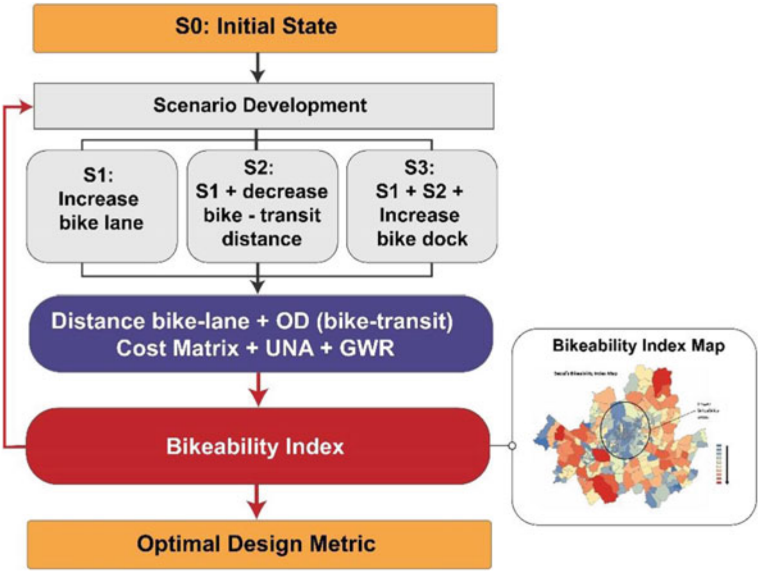


Fig. 9.3 Bike planning support system flow chart

routes and choosing the best routes (Makarova et al. 2019). Their objectives were to provide both strategic and operational capabilities by improving environmental efficiency, safety, and passenger traffic, and reducing the last mile (Makarova et al. 2019). Unlike Makarova et al., our research's PSS (Fig. 9.3) is not focused on improving safety, which could be included in further studies. But similarly, our PSS enables the investigation of several factors to increase the bikeability of a district.

We developed three scenarios. To develop adequate scenarios, we investigated the local influence of each explanatory parameter and drafted intervention measures accordingly (Fig. 9.4). The implemented PSS scenarios were: (1) increasing the bike-only lanes, (2) increasing the bike lane while decreasing the distance between bike and transit stations, and (3) increasing the dock number and bike lanes while decreasing bike–transit distance (Fig. 9.5). For each scenario, the UNA centrality indices were computed and input into the bikeability index model. The calculations of the bikeability index continued in new iterations until we reached the optimal design metric that increased the bikeability index for the entire city (Fig. 9.3). This optimal design metric solution was designed to increase the efficacy of interventions while improving the performance of the bike system.

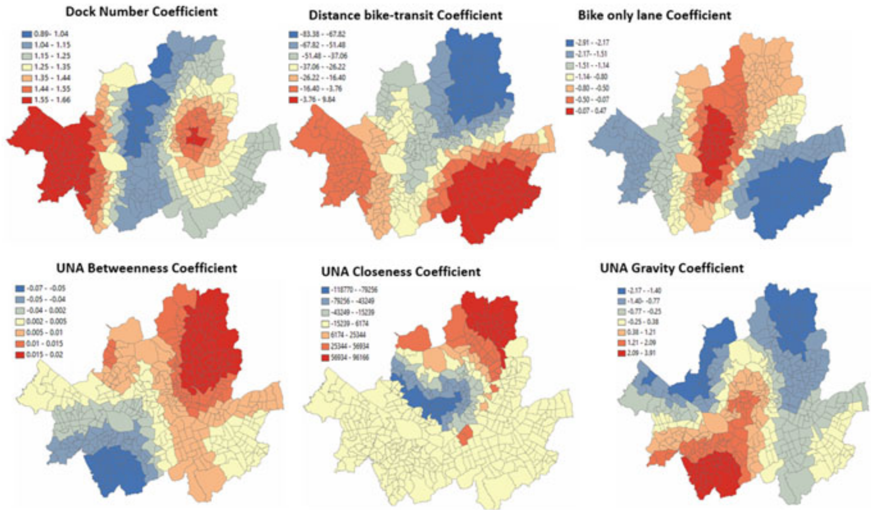


Fig. 9.4 GWR local influence of explanatory variables

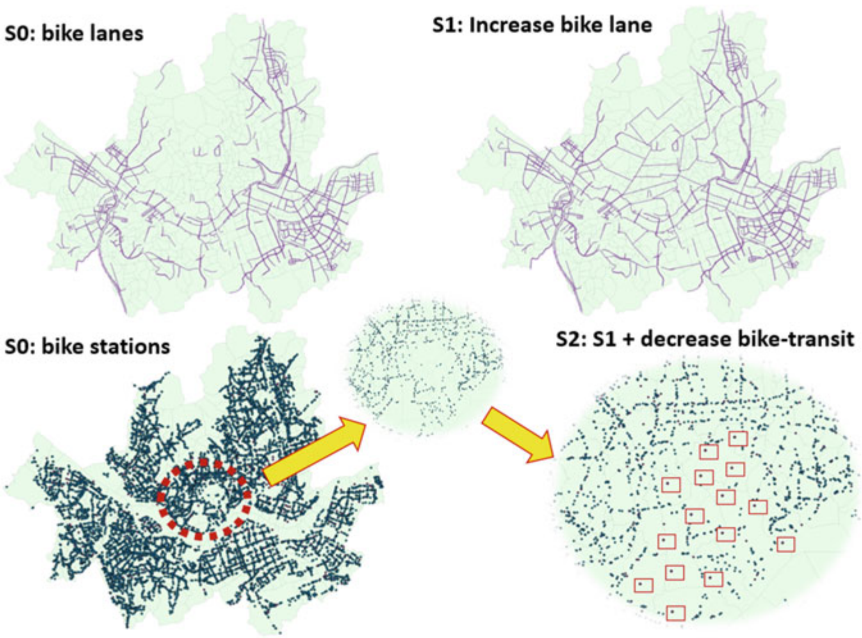


Fig. 9.5 Bike-developed scenarios for the PSS

9.3.2 *Development of the Model's Dataset*

Most of the data were downloaded from the Seoul metropolitan government's open data website. Data included socioeconomics (population and employment data), built and natural environment (incline and land-use data), station locations (bike, bus, and subway), usage demand for each station (bike, bus, and subway), and bike dock numbers by station. The retail consumption was extracted from credit card data. However, to prepare the input dataset, we needed to compute the following factors: slope index (degree of incline) per district; land-use mix index per district using the entropy theory; the bike–transit distance using the original-destination (OD) cost matrix, considering transit stations as origins and bike stations as destinations; and centrality indices using the MIT UNA tool.

9.3.2.1 Land-Use Mix Index and Slope Index

Many city-development branches, such as TOD, apply land-use mix theories. There are two popular methods to measure the land-use mix: the entropy index (EI), which considers the relative percentage of existing land-use types within an area; and the Herfindahl–Hirschman index (HHI), which reflects the level of the land-use mix of a specific area (Zagorskas 2016).

For this study we adopted the EI index because our focus was on the relative percentage of existing land-use types. Seoul's land-use data contain three hierarchical levels, and we considered only the first level. The land-use mix index was calculated according to each district in Seoul (Fig. 9.6).

To determine slope, we downloaded raster data files and transformed them into vector data before splitting them based on the administrative districts. The adopted slope unit was the degree, and these values were divided into nine groups. The slope index (Fig. 9.6) was calculated for each district.

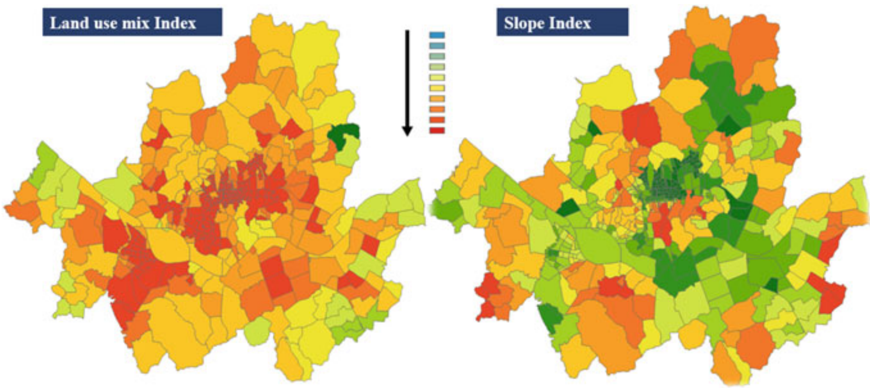


Fig. 9.6 Land-use mix and slope indices

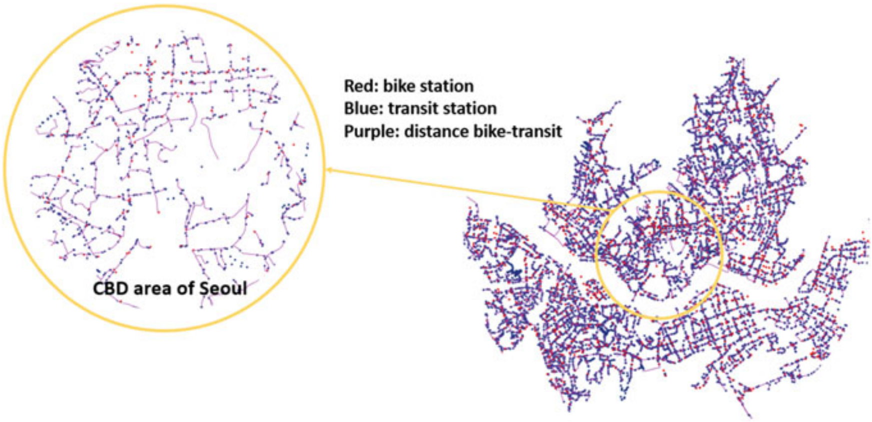


Fig. 9.7 Bike-transit distance (CBD zoom)

9.3.2.2 Distance Between Bike and Public Transit

We calculated an origin–destination cost matrix using ArcGIS to reflect the impact of the proximity of a transit station in the analysis (Fig. 9.7). As shown in Fig. 9.7, the study calculated the shortest path from a bike station to transit station, with a zoom on the Central Business District (CBD) area. The distance was aggregated on a district level using the mean value. We considered both rail and bus stations in the analysis.

9.3.2.3 Centrality Indices Measurement

A transport mode's accessibility has been linked to its centrality indices; these indices capture the importance of station locations within a road network (Rubulotta et al. 2013; Sarlas et al. 2020). Centrality parameters describe urban locations (Wang et al. 2011), capturing their significance (nodes or links) within a network based on different criteria, either geometrical or topological (Agryzkov et al. 2014; Rubulotta et al. 2013). Thus, our study calculated the centrality indices for bike stations to reflect their accessibility and to identify which centrality index must be considered when planning bike stations. UNA centrality indices are reach, gravity, betweenness, closeness, and straightness. The UNA toolbox was used to compute these centrality indices within a search radius of 800 m, which was the access shed identified for walkability to a public transit station (El-Geneidy et al. 2014). Then, the centrality metrics were aggregated at the district level (Fig. 9.8).

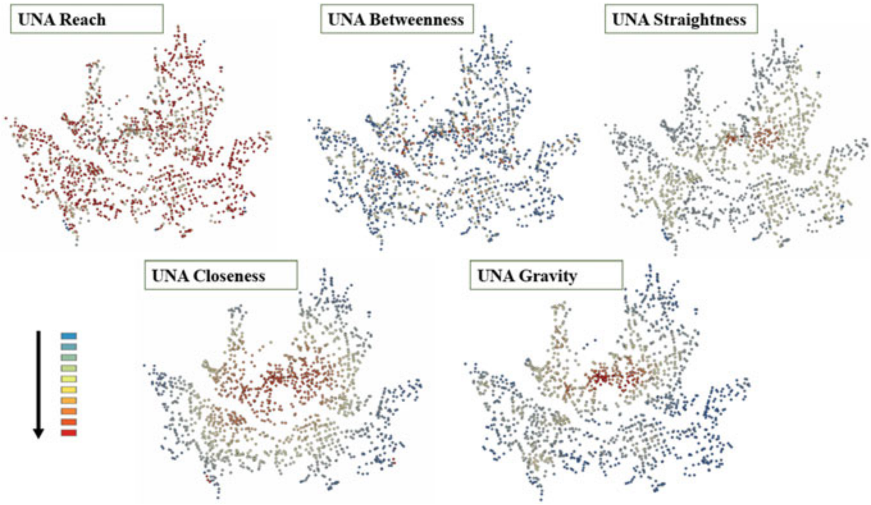


Fig. 9.8 UNA indices for bike stations

9.3.3 Development of the Models: Poisson and GWR

We used the Poisson regression model because cycling data are in counts (Orvin et al. 2021). The Poisson model permitted the removal of variables with high multicollinearity and the identification of a preliminary model that included key variables. However, before developing a GWR model, the spatial autocorrelation, dependence, and heterogeneity need to be verified using the Moran index, Lagrange multiplier, and Chow test, respectively (Burrows and Cantrell 1990; Li et al. 2007; Longley and Tobón 2004). The GWR is valid when these three spatial criteria are proved.

Different explanatory factors might operate at different spatial scales, resulting in coefficients with varying spatial heterogeneity levels (Chen and Mei 2021). According to Chen and Mei (2021), the GWR model is expressed by Eq. (9.1), the spatially varying coefficients are estimated using the locally weighted least-squares method, and the estimated coefficients provide information regarding the spatial heterogeneity of the regression relationship.

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{r} \beta_j(u_i, v_i)x_{ij} + \varepsilon_i, i = 1, 2, \ldots, n \tag{9.1}$$

where,

$\{y_i; x_{i1}, x_{i2}, \ldots, x_{ir}\}$ with $i = \{1, \ldots, n\}$ are the observations of response variable Y and explanatory variables X_1, X_2, …, X_r at n spatial sampling locations,

$\{(u_i, v_i)\}$ with $i = \{1, \ldots, n\}$ are the geographic coordinate of the spatial locations,

$\{\beta_j(u, v)$ with $j = \{0, \ldots, r\}$ are the $r + 1$ unknown spatially varying coefficient to be estimated.

$\{\varepsilon_i\}$ $i = \{1, \ldots, n\}$ are independent and identically distributed errors, with $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma 2$.

9.4 Results and Discussion

9.4.1 Centrality Indices Using UNA Tool

In this research we conducted an urban network analysis (UNA) to compute the five centrality indices for bike stations in Seoul (Fig. 9.6). The initial multicollinearity test showed that the "reach" and "straightness" measures had a high value for the variance inflation factor (VIF), which led to the exclusion of the two measures in the following analyses. Therefore, we only considered "betweenness," "closeness," and "gravity" in further analyses. These three measures were computed for each scenario and each iteration that were then used in the bikeability index calculation. Results showed that the second scenario was the best (Fig. 9.9) because it improved all three UNA measures. The second scenario consisted of increasing both bike-lane and bike-dock numbers while decreasing the distance to transit either by creating new stations near key rail or bus stations or by relocating some bike stations closer to key rail or bus stations.

9.4.2 Spatial Validation Tests: Autocorrelation, Dependence, and Heterogeneity

In conducting spatial validation tests, we first performed a multicollinearity test. Variables with VIF scores higher than 5 were excluded. After removing the predictors with the highest VIF scores, the new Poisson regression revealed an R2 value of 82%. The selected variables were land-use mix index, slope, population and employment density, transit demand, number of transit stations, dock number, bike–transit distance, and three centrality indices (closeness, betweenness, and gravity).

Using the Poisson residuals table, we ran the Moran's I statistical test to check the global and local spatial autocorrelation. The results showed less than a 1% likelihood that the clustered pattern could be a result of random chance. Therefore, the Moran's index (9.14) was statistically significant with an almost null *p*-value. The results of the Lagrange multiplier proved the existence of spatial dependence. The test revealed a minimal *p*-value of 0.0004 for the spatial error model, whereas the *p*-value of the Lag model was about 0.05. Thus, the research proved the existence of spatial dependence.

The Chow test measures heterogeneity based on structural stability testing using the residual sum of squares in constrained and unconstrained regressions. To run

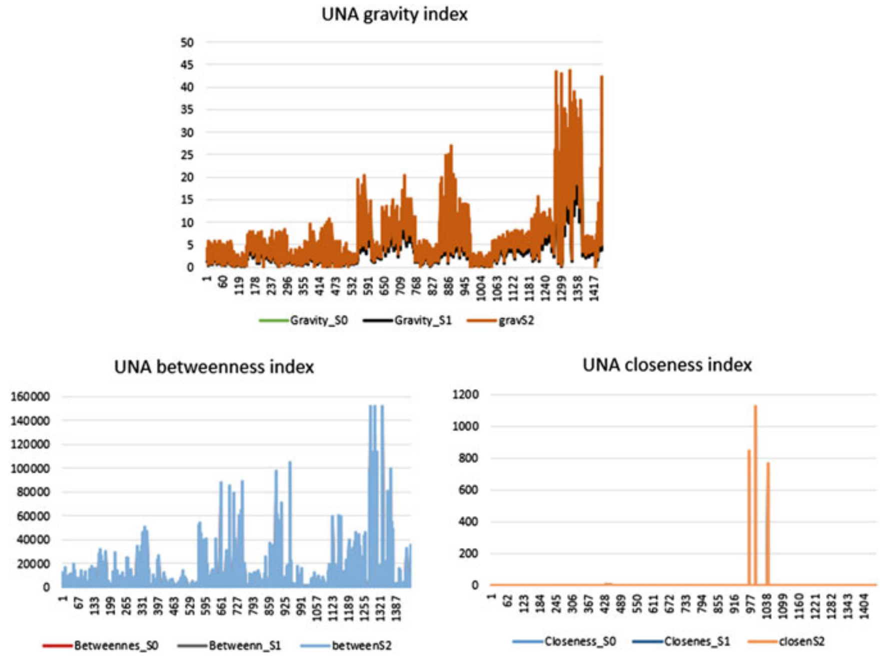


Fig. 9.9 Centrality indices results using the UNA tool

the Chow test, we first ran a regression of the entire dataset (pooled regression), then we randomly split the dataset into two groups based on the 50:50 rule and ran two separate regressions. The results revealed that the F-value was 18.0797, and the *p*-value was almost 0 (4.8e-17). Because the *p*-value was small and the F-statistics exceeded the critical F-value, we rejected the null hypothesis that the two regressions of the dataset groups were equal. Therefore, this study demonstrated the existence of spatial heterogeneity.

The existence of spatial autocorrelation, dependence, and heterogeneity validated the implementation of the GWR model to investigate the existing spatial relationships and derive a bikeability index for the city of Seoul (Fig. 9.10).

9.4.3 Geographically Weighted Regression Model

The previous spatial tests (Fig. 9.10) proved the validity of the application of the GWR model. Therefore, we used the GWR model to investigate relationships across geographic spaces, clarifying how the locally weighted regression coefficients deviated from their globally weighted counterparts. Thus, GWR permitted the identification of local variables that influence bike demand within each district in Seoul. The GWR uses an adaptive kernel-type technique to reflect the feature distribution

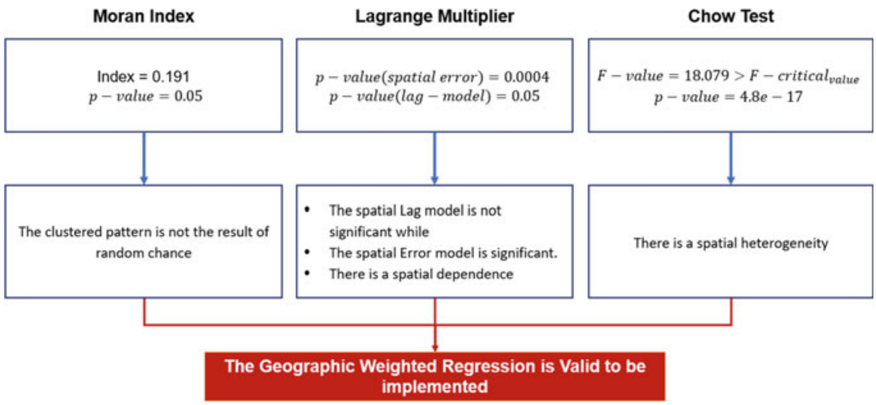


Fig. 9.10 Spatial tests for the validity of GWR

in the analysis—the spatial context in this case—as it is represented by a function of a specific number of neighbors. The corrected Akaike information criterion (AIC) was used for the bandwidth method.

The GWR model showed high goodness of fit for several reasons. First, only a few districts exhibited a slightly higher residual standard deviation (Fig. 9.11). Second, the condition number was low and under 30, with a maximum of 25.25 within very few districts (Fig. 9.12). And third, high local R^2 values (higher than 0.5) were exhibited in all districts (Fig. 9.12). When comparing the Poisson and GWR models, we found that the GWR revealed better performance and reliability with a higher R^2 value (Table 9.1).

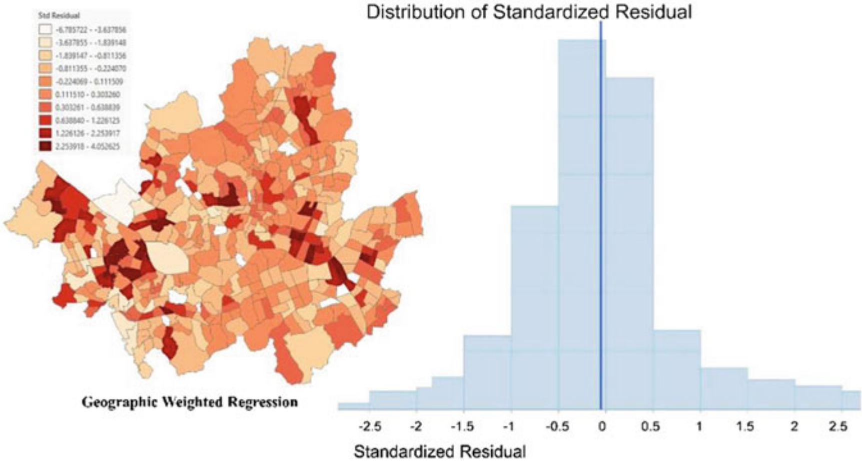


Fig. 9.11 GWR standard residual map

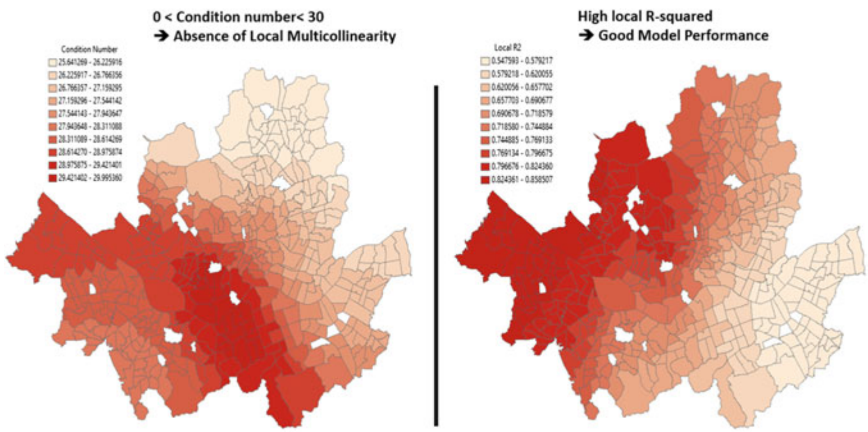


Fig. 9.12 GWR model's performance

Table 9.1 Results of the two models (GWR and poisson)

Evaluation factor	GWR model	Poisson model
R-square (%)	83.2	82.7
AICc	4,465.41	8,931.35

9.5 Bikeability Index-Based Planning Support System

As a result of the GWR displaying high accuracy, performance, and goodness of fit, we selected the model to represent the bikeability index for the city of Seoul and used the model to generate a bikeability index map (Fig. 9.13 and Eq. 9.2). The index comprised the following independent variables: employment and population densities, land-use mix and slope indices, retail consumption, centrality indices (betweenness, closeness, and gravity), dock number, bike–transit distance, and transit and bike demand.

$$BI_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{k,i} + \epsilon_i \tag{9.2}$$

where

i represents a district id,
BI_i is the bikeability index of district i,
ϵ_i is the error at district i,
(u_i, v_i) represents coordinates x and y of district i,
$\beta_0(u_i, v_i)$ is the intercept parameter of district i, and
$\beta_k(u_i, v_i)$ is the coefficient of the predictor variable x_k at district i.

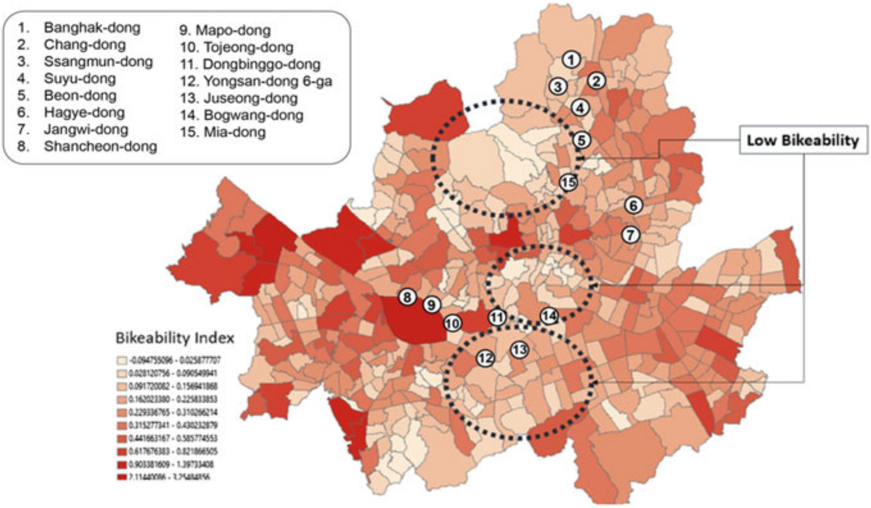


Fig. 9.13 Bikeability index map of Seoul

A comparison of the bicycle observed demand and the predicted demand (Fig. 9.14), using the bikeability index, indicated that the shared-bike program in Seoul is not inefficient. However, results identified many areas that could achieve higher efficacy and performance. These areas displayed a significant difference between the bikeability index value and the real bike demand, suggesting that the bike demand potential is higher than the actual demand (Fig. 9.14). For instance, Yongsan-dong 6-ga, Bogwang-dong, Mapo-dong, Shancheon-dong, Tojeong-dong, Bogwang-dong, Dongbinggo-dong, and Juseong-dong have high bikeability indices and thus show high biking potential but low demand (Figs. 9.13 and 9.14). Furthermore, the Nohyeon-dong, Sinsa-dong, Jamsil-dong, Samjeon-dong, Rwon-dong, and Gaepo-dong areas demonstrated that they might achieve higher demand because of their high bikeability indices (Figs. 9.13 and 9.14). The latter districts are expected to increase active transportation because the Seoul government is creating an exclusive lane for bikes and personal mobility scooters. In addition, the Seoul government is implementing congestion charges and low-emission zone restrictions within the latter areas. If accompanied by a biking policy, these measures are expected to help districts reach their full biking potential. In contrast, some areas such as Banghak-dong, Ssangmun-dong, Chang-dong, Suyu-dong, Beon-dong, Mia-dong, Hagye-dong, and Jangwi-dong show over-performance, with higher-than-expected biking demand (Figs. 9.13 and 9.14). Such areas require in-depth analysis to identify the factors that help people become more engaged in biking activities and examine the findings for other areas with high potential but low demand. For instance, recent strategies to improve the performance of cycling transportation systems are based on infrastructural, behavioral, and multimodal measures, which are strongly tied to bikeability (Castañon and Ribeiro 2021). A demand-forecasting study on a cycle

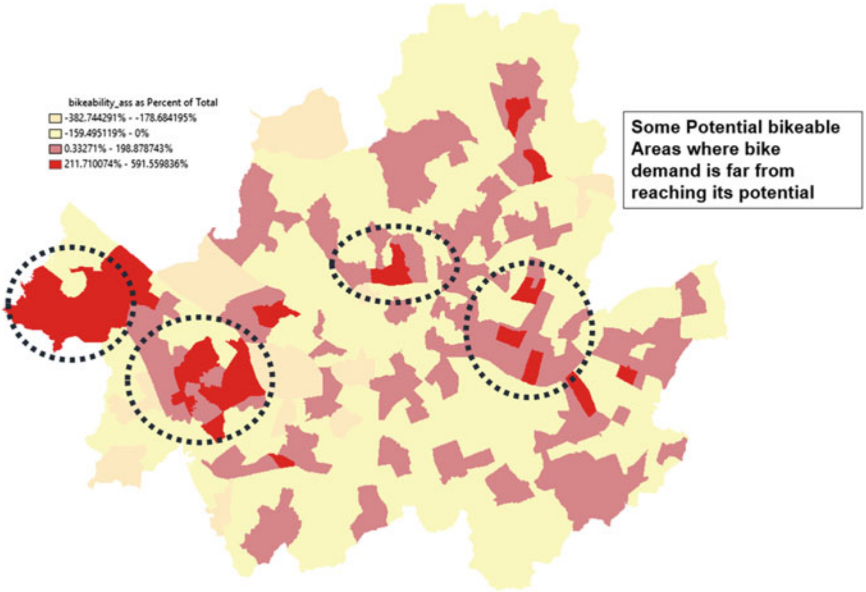


Fig. 9.14 Comparing predicted and observed demand

superhighway scheme in Seoul found that implementing an exclusive cycling lane had a significant impact, with 30% of short-trip passengers converting to cycling modes of transportation (Ku et al. 2021). An investigation of the impacts of the coronavirus disease 2019 (COVID-19), weather conditions, and social index on the Seoul shared-bike system revealed that outbreaks of COVID-19 had different impacts on shared-bike usage based on spatial characteristics and that retail and leisure activities exhibited significant influence (Kwak et al. 2021).

Further analysis of each variable's area of influence along with the analysis of the values and distribution of the explanatory variables themselves would provide a reliable understanding of the bikeability and the specific characteristics that influence biking within each district. This understanding would provide basic input data for any future biking policy. For instance, within Area-1, the dock and transit stations number displayed low impact, while the land-use mix and gravity indices displayed higher local influence (Fig. 9.15). Therefore, to increase biking within Area-1, it is recommended to increase the land-use mix and gravity indices to have the most significant impact on enhancing bikeability.

The GWR showed spatial heterogeneity in the relationship between biking and the certain variables within the city. Therefore, in many districts in the study a deeper investigation could be conducted based on this index, to extract further information and understanding. For example, using the bikeability index as an assessment tool in targeted areas, we found that certain areas could be improved in terms of the number of docks, the transit distance, and bike station gravity index; the assessment tool showed that applying these interventions would result in significant improvement

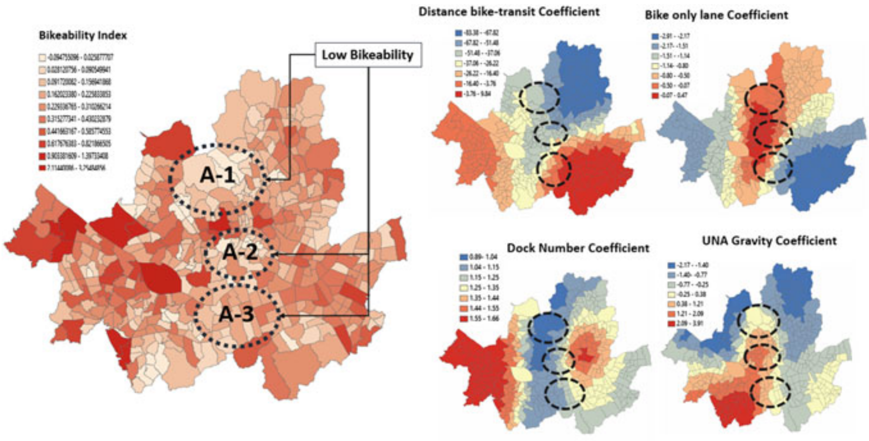


Fig. 9.15 Bikeability index as an assessment tool

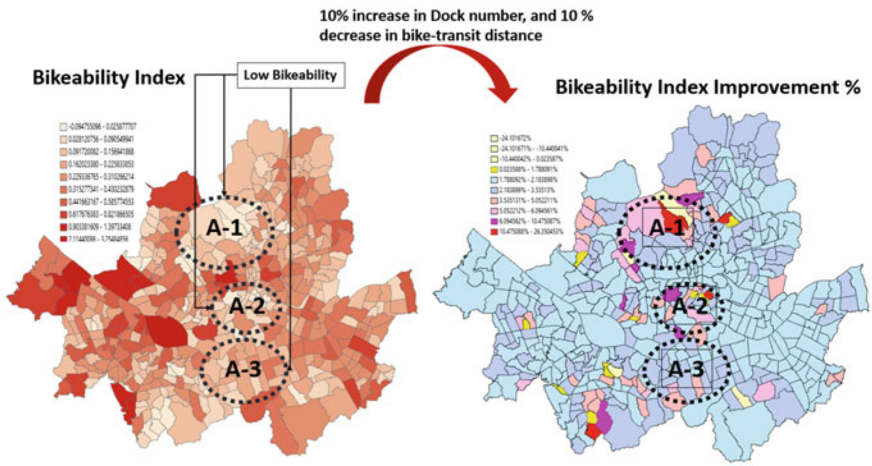


Fig. 9.16 Targeted bikeability improvement tool

in the bikeability index in these targeted areas (Fig. 9.16). However, this study was designed to improve the entire city's bikeability; thus, three scenarios (Fig. 9.3) were adopted all over the city with consideration for personalized improvement strategies according to the local and global impact of the studied variables.

Results showed that the bikeability index of the second scenario (increasing cycling-only lanes) was the best compared to the other two scenarios (adding bicycle docking stations, and adding lanes plus decreasing bike–transit distance) in all iterations (Fig. 9.17). Thus, aggressive interventions that include many significant measures are much preferred to achieve better results. The level of implementation of each scenario depends on the characteristics of each area, but generally the increase

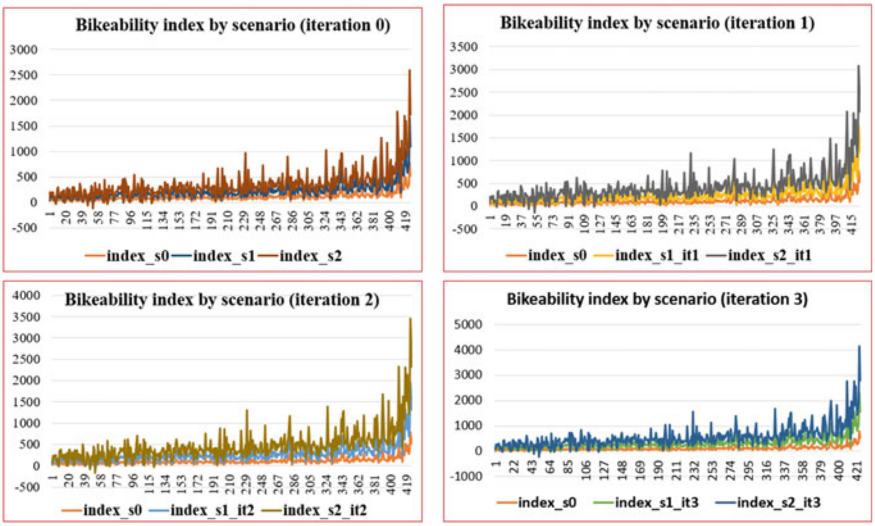


Fig. 9.17 Bikeability index by scenario and iteration

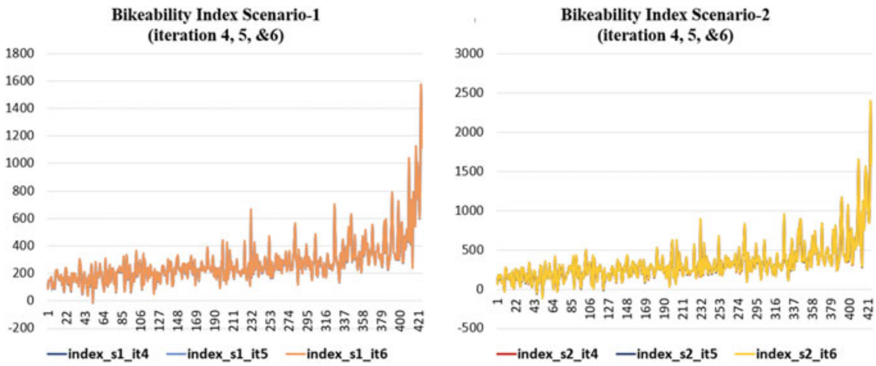


Fig. 9.18 Bike PSS's optimal design metric

in bicycle lanes and docks was about 10% for the first iteration, 15% for the second, 22% for the third, and 26% for the fourth. The distance to public transit was focused on areas with high bikeability potential and low actual demand. Improvements would consist of adding bike docking stations within areas of high economic activity and relocating bike stations within areas of the highest demand for their use.

Moreover, results showed that after the fourth iteration, improvements in the bikeability index were small (Fig. 9.18). Thus, conducting further interventions would not be cost-effective. Therefore, the optimal design metric solution was achieved at the fourth iteration for all scenarios (Fig. 9.18).

9.6 Conclusion and Policy Implications

Seoul's government is planning to build an eco-friendly city by focusing on improving active transportation modes, such as cycling and personal mobility. The government has implemented ambitious projects for shared bikes and personal mobility, and these efforts have shown a significant increase in the bicycle usage rate but have not reached the expected efficiency. The pressure to reach higher efficiency is increasing given the amount of money invested in these projects. Therefore, in this study we developed a bikeability index for the city of Seoul to measure bikeability in each district based on socioeconomic and built environment data. The goal was to find areas to increase efficacy in the shared bicycle programs.

The results showed that the number of docking stations, geographic inclines (slope), centrality indices (betweenness, closeness, gravity), economic activity, population, and distance from docking to public transit (bike–transit distance) are the most significant variables in determining the bikeability of each district. The geographically weighted regression (GWR) analysis showed good accuracy and performance in computing the bikeability index in each district. The index showed some areas of anomaly when comparing the predicted and the observed demand. The model permits investigation of the global and local impact of each explanatory variable, thereby facilitating the design of specific policy interventions to increase efficiency in different districts according to their distinct characteristics.

In summary, this study is the first to investigate the bikeability of Seoul through the implementation of a bike planning support system (PSS). The PSS developed in this study employs a comprehensive dataset that includes the socioeconomic data (population and employment), built and natural environment (slope and land-use mix), public transit and bike demand and infrastructure (station locations and number, and bike–transit distance), economic activity (retail consumption), and finally centrality indices to investigate the bikeability by the district. However, the large scale of each district and variability in the availability of data impacted the accuracy of the index. Therefore, for future studies, we would advise implementing a smaller-scale analysis. Additional socioeconomic data (e.g., income and housing type), safety data, and spatial data (e.g., green parks and leisure facilities) would enhance the bikeability index's results. Users' perceived satisfaction with the shared-bike system would be powerful additional data to enhance the developed bike PSS much further.

The bike PSS developed in this study could be considered a stepping-stone toward zero-carbon-oriented transportation policies, by implementing components related to public transportation and mobility as a service (MaaS). Therefore, we drafted the following recommendations:

(1) Minimize analysis scale to deepen the understanding and develop better interventions.
(2) Employ this bike PSS to achieve more eco-friendly cycling planning.
(3) Enhance cycling planning by employing additional surveys, especially within anomalous areas, to ensure the implementation of adequate and efficient cycling policy.

References

Agryzkov T, Oliver JL, Tortosa L, Vicent J (2014) A new betweenness centrality measure based on an algorithm for ranking the nodes of a network. Appl Math Comput 244:467–478. https://doi.org/10.1016/j.amc.2014.07.026

Arellana J, Saltarín M, Larrañaga AM, González VI, Henao CA (2020) Developing an urban bikeability index for different types of cyclists as a tool to prioritise bicycle infrastructure investments. Transp Res Part A Policy Pract 139:310–334. https://doi.org/10.1016/j.tra.2020.07.010

Burrows PM, Cantrell RS (1990) Specification errors and the chow test. Econ Lett 34(2):131–135. https://doi.org/10.1016/0165-1765(90)90232-P

Caggiani L, Ottomanelli M (2012) A modular soft computing based method for vehicles repositioning in bike-sharing systems. Procedia Soc Behav Sci 54:675–684. https://doi.org/10.1016/j.sbspro.2012.09.785

Castañon UN, Ribeiro PJG (2021) Bikeability and emerging phenomena in cycling: exploratory analysis and review. Sustainability 13(4):2394. https://doi.org/10.3390/su13042394

Cervero R, Kockelman K (1997) Travel demand and the 3Ds: density, diversity, and design. Transp Res Part D: Transp Environ 2(3):199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Chen F, Mei C-L (2021) Scale-adaptive estimation of mixed geographically weighted regression models. Econ Model 94:737–747. https://doi.org/10.1016/j.econmod.2020.02.015

El-Geneidy A, Grimsrud M, Wasfi R, Tétreault P, Surprenant-Legault J (2014) New evidence on walking distances to transit stops: identifying redundancies and gaps using variable service areas. Transportation 41(1):193–210. https://doi.org/10.1007/s11116-013-9508-z

Eren E, Uz VE (2020) A review on bike-sharing: the factors affecting bike-sharing demand. Sustain Cities Soc 54:101882. https://doi.org/10.1016/j.scs.2019.101882

Fikar C, Hirsch P, Gronalt M (2018) A decision support system to investigate dynamic last-mile distribution facilitating cargo-bikes. Int J Logist Res Appl 21:300–317. https://doi.org/10.1080/13675567.2017.1395830

Geertman S, Stillwell J (eds) (2003). Springer, Berlin Heidelberg, Berlin, Heidelberg

Getis A (2007) Reflections on spatial autocorrelation. Reg Sci Urban Econ 37(4):491–496. https://doi.org/10.1016/j.regsciurbeco.2007.04.005

Grigore E, Garrick N, Fuhrer R, Axhausen IKW (2019) Bikeability in Basel. Transportation Research Record: Journal of the Transportation Research Board 2ss673(6):607–617. https://doi.org/10.1177/0361198119839982

Hankey S, Lindsey G, Marshall JD (2017) Population-level exposure to particulate air pollution during active travel: planning for low-exposure, health-promoting cities. Environ Health Perspect 125(4):527–534. https://doi.org/10.1289/EHP442

Hardinghaus M, Nieland S, Lehne M, Weschke J (2021) More than bike lanes—a multifactorial index of urban bikeability. Sustainability 13(21):11584. https://doi.org/10.3390/su132111584

Jiao J, Lee HK, Choi SJ (2022) Impacts of COVID-19 on bike-sharing usages in Seoul South Korea. Cities 130:103849. https://doi.org/10.1016/j.cities.2022.103849

Kamel MB, Sayed T, Bigazzi A (2020) A composite zonal index for biking attractiveness and safety. Accid Anal Prev 137:105439. https://doi.org/10.1016/j.aap.2020.105439

Kellstedt DK, Spengler JO, Foster M, Lee C, Maddock JE (2021) A scoping review of bikeability assessment methods. J Community Health 46(1):211–224. https://doi.org/10.1007/s10900-020-00846-4

Ku D, Kwak J, Na S, Lee S, Lee S (2021) Impact assessment on cycle super highway schemes. Chem Eng Trans 83(February):181–186. https://doi.org/10.3303/CET2183031

Kwak J, Oh H, Jeong I, Shin S, Ku D, Lee S (2021) Changes in shared bicycle usage by COVID-19. Chem Eng Trans 89(April):169–174. https://doi.org/10.3303/CET2189029

Lee S, Bencekri M (2021) Urban form and public transport design. Urban Form and Accessibility. pp 289–306. https://doi.org/10.1016/b978-0-12-819822-3.00018-3

Li H, Calder CA, Cressie N (2007) Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model. Geogr Anal 39(4):357–375. https://doi.org/10.1111/j.1538-4632.2007.00708.x

Lowry MB, Callister D, Gresham M, Moore B (2012) Assessment of communitywide bikeability with bicycle level of service. Transp Res Rec J Transp Res Board 2314(1):41–48. https://doi.org/10.3141/2314-06

Lu Y, Yang Y, Sun G, Gou Z (2019) Associations between overhead-view and eye-level urban greenness and cycling behaviors. Cities 88:10–18. https://doi.org/10.1016/j.cities.2019.01.003

Makarova I, Boyko A, Tsybunov E, et al (2019) The use of the decision support system to control bicycle transportation. In: Proceedings of the 5th international conference on vehicle technology and intelligent transport systems. SCITEPRESS - Science and Technology Publications, pp 649–656

Mrkajić V, Anguelovski I (2016) Planning for sustainable mobility in transition cities: cycling losses and hopes of revival in Novi Sad, Serbia. Cities 52:66–78. https://doi.org/10.1016/j.cities.2015.11.029

Munira S, Sener IN, Zhang Y (2021) Estimating bicycle demand in the Austin, Texas area: Role of a bikeability index. J Urban Plan Dev 147(3). https://doi.org/10.1061/(ASCE)UP.1943-5444.0000725

Organisation for Economic Co-operation and Development (OECD). (2021). Korea 1. https://www.itf-oecd.org/sites/default/files/korea-road-safety.pdf. Accessed 1 May 2020

Orvin MM, Fatmi MR, Chowdhury S (2021) Taking another look at cycling demand modeling: a comparison between two cities in Canada and New Zealand. J Transp Geogr 97:103220. https://doi.org/10.1016/j.jtrangeo.2021.103220

Pettit C, Bakelmun A, Lieske SN et al (2018) Planning support systems for smart cities. City Cult Soc 12:13–24. https://doi.org/10.1016/j.ccs.2017.10.002

Porter AK, Kohl HW, Pérez A, Reininger B, Pettee Gabriel K, Salvo D (2020) Bikeability: assessing the objectively measured environment in relation to recreation and transportation bicycling. Environ Behav 52(8):861–894. https://doi.org/10.1177/0013916518825289

Rubulotta E, Ignaccolo M, Inturri G, Rofè Y (2013) Accessibility and centrality for sustainable mobility: regional planning case study. J Urban Plan Dev 139(2):115–132. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000140

Sarlas G, Páez A, Axhausen KW (2020) Betweenness-accessibility: estimating impacts of accessibility on networks. J Transp Geogr 84:102680. https://doi.org/10.1016/j.jtrangeo.2020.102680

Schmid-Querg J, Keler A, Grigoropoulos G (2021) The Munich bikeability index: a practical approach for measuring urban bikeability. Sustainability 13(1):428. https://doi.org/10.3390/su13010428

Sevtsuk A, Mekonnen M (2012) Urban network analysis: a new toolbox for measuring city form in ArcGIS. Simul Ser 44(8 Book):133–145

Shaer A, Rezaei M, Moghani Rahimi B, Shaer F (2021) Examining the associations between perceived built environment and active travel, before and after the COVID-19 outbreak in Shiraz City, Iran. Cities 115:103255. https://doi.org/10.1016/j.cities.2021.103255

Shin H, Cagnina C, Basiri A (2022) The impact of built environment on bike commuting: utilising strava bike data and geographically weighted models. AGILE: GIScience Ser 3:1–9. https://doi.org/10.5194/agile-giss-3-15-2022

te brömmelstroet 2009 Te Brömmelstroet M (2009) The relevance of research in planning support systems: a response to Janssen et al. Environ Plan B Plan Des 36:4–7. https://doi.org/10.1068/b3601com

Vidal Tortosa E, Lovelace R, Heinen E, Mann RP (2021) Cycling behaviour and socioeconomic disadvantage: an investigation based on the english national travel survey. Transp Res Part A: Policy Pract 152:173–185. https://doi.org/10.1016/j.tra.2021.08.004

Vonk G, Geertman S, Schot P (2007) A SWOT analysis of planning support systems. Environ Plan A Econ Sp 39:1699–1714. https://doi.org/10.1068/a38262

Wang J, Mo H, Wang F, Jin F (2011) Exploring the network structure and nodal centrality of China's air transport network: a complex network approach. J Transp Geogr 19(4):712–721. https://doi.org/10.1016/j.jtrangeo.2010.08.012

Wei Z, Zhen F, Mo H, Wei S, Peng D, Zhang Y (2021) Travel behaviours of sharing bicycles in the central urban area based on geographically weighted regression: the case of Guangzhou, China. Chin Geogr Sci 31(1):54–69. https://doi.org/10.1007/s11769-020-1159-3

Winters M, Brauer M, Setton EM, Teschke K (2013) Mapping bikeability: a spatial tool to support sustainable travel. Environ Plann B Plann Des 40(5):865–883. https://doi.org/10.1068/b38185

Xin R, Ai T, Ding L, Zhu R, Meng L (2022) Impact of the COVID-19 pandemic on urban human mobility—a multiscale geospatial network analysis using New York bike-sharing data. Cities 126:103677. https://doi.org/10.1016/j.cities.2022.103677

Yang Y, Wu X, Zhou P, Gou Z, Lu Y (2019) Towards a cycling-friendly city: an updated review of the associations between built environment and cycling behaviors (2007–2017). J Transp Health 14:100613. https://doi.org/10.1016/j.jth.2019.100613

Zagorskas J (2016) GIS-based modelling and estimation of land use mix in urban environment 2: problem formulation. Int J Environ Sci 1:284–293

Zayed MA (2016) Towards an index of city readiness for cycling. Int J Transp Sci Technol 5(3):210–225. https://doi.org/10.1016/j.ijtst.2017.01.002

Zhang Y, Thomas T, Brussel M, van Maarseveen M (2017) Exploring the impact of built environment factors on the use of public bikes at bike stations: case study in Zhongshan, China. J Transp Geogr 58:59–70. https://doi.org/10.1016/j.jtrangeo.2016.11.014

Zuo T, Wei H, Chen N, Zhang C (2020) First-and-last mile solution via bicycling to improving transit accessibility and advancing transportation equity. Cities 99:102614. https://doi.org/10.1016/j.cities.2020.102614

Chapter 10
Integrating Big Data and a Travel Survey to Understand the Gender Gap in Ride-Hailing Usage: Evidence from Chengdu, China

Si Qiao, Anthony Gar-On Yeh, and Mengzhu Zhang

Abstract Improving transport systems to increase women's access to social opportunities and essential facilities has been key to reducing gender inequality. Studies have examined the gendered nature of travel from the perspective of a mismatch between women's needs and availability of transport services, including fragmentized activity space, low affordability, and sensitization to safety. However, minimal attention has been given to the gender gap in the age of ride-hailing. Thus, this paper examines the nexus between gender and ride-hailing usage from the aspect of activity space and affordability. Two key questions are explored: (a) Are women dependent on ride-hailing? (b) If ride-hailing serves women differentially, how does this gender difference in the use of ride-hailing services occur? An innovative integration of big data and a travel survey is developed to examine such questions in Chengdu, China. Survey results and modelling analysis indicate that gender gaps in mobility is relatively mitigated by ride-hailing.

Keywords Gender gap · Social inequality · Platform-based mobility service · Activity space · Transport equity

S. Qiao · A. G.-O. Yeh (✉)
Department of Urban Planning and Design, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China
e-mail: hdxugoy@hkucc.hku.hk

S. Qiao
e-mail: siqiao@connect.hku.hk

M. Zhang
School of Urban Planning and Design, Shenzhen Graduate School, Peking University, Shenzhen, China
e-mail: mengzhu.zhang@pku.edu.cn

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_10

10.1 Introduction

Smart mobility, or a platform-based on-demand mobility service, generally refers to a new travel mode in which users can obtain car use rights or mobility service on demand through smartphones and pay electronically, and has received wide attention (Clewlow and Mishra 2017; Contreras and Paz 2018; Young and Faber 2019). This concept covers a wide range of travel choices such as ride-hailing/ridesourcing, ridesplitting, and e-hail (Shaheen and Chan 2016). Among them, the most popular and leading trend that has reshaped todays urban transport system is the ride-hailing mode (*WangYueChe*, in Chinese), which is provided by several transport network companies (TNCs) such as Uber, Lyft, and DiDi.

The advantages of using ride-hailing mode include benefits to urban efficiency, society and the environment such as reduced congestion, carbon emissions, vehicle usage, private car ownership, and higher vehicle miles or kilometers (Shaheen 2018; Shaheen and Chan 2016; Shaheen et al. 2015). Moreover, ride-hailing offers opportunities to expand access beyond the limitations of traditional transport infrastructure by solving the first/last mile problem (Shaheen and Chan 2016; Shaheen 2018; Jin et al. 2018; Ward et al. 2019).

On the other hand, an increasing number of scholars have expressed concerns with unequal access to ride-hailing services among social groups, which may cause the (re)production of transport poverty and social exclusion (Ge et al. 2016; Rayle et al. 2016; Deka and Fei 2019; Zhang et al. 2020; Qiao and Yeh 2021). It is hypothesized that ride-hailing underprivileges certain populations with respect to transportation choices. For instance, "Uberization" attracts transit-dependent individuals to ride-hailing (Kong et al. 2020), thereby resulting in a decline in transit ridership, which in turn accelerates cuts to transit service (Clewlow and Mishra 2017; Rayle et al. 2016). Eventually, travel choices for disadvantaged populations, such as the elderly, the poor, and housewives, may be reduced, thereby limiting their access to job opportunities, healthcare, shopping, and other social activities.

Although the results of the association between individual/neighborhood characteristics and ride-hailing accessibility have remained mixed, several scholars have been unanimous about social and geographical inequalities in ride-hailing. Therefore, understanding existing ride-hailing applications and their potential impact on such disadvantaged groups is imperative for planners and policymakers in formulating strategies for future transport development. However, gender, as a key dimension in which social/transport inequality is (re)produced, has not received sufficient attention.

For women, a group that comprises nearly half of the population, ride-hailing usage may be able to change long-standing inequalities in transportation access if policy makers could better understand it. According to the World Bank, 62 percent of women (aged above 15 years old) in China participate in the labor market. Although the labor force participation of women is high, they still do most of the caregiving for their families, such as childcare-related duties, buying ingredients for cooking, and housekeeping. Such activities trap them in a restricted activity space, limit their daily mobility, and thus force them to seek jobs near their homes (Kwan 1999). Short and fragmented travel needs are difficult to accommodate in public transport designed for connecting work-related hotspots. Moreover, carrying heavy paraphernalia and caring for small children are constant struggle for mothers when they travel. They also have less car access than men. The common reasons are twofold. Men-oriented allocation of household cars (Scheiner and Holz-Rau 2012; Solá 2016) and stigmatization against women driving (Solá 2016). Such factors prevent women from using automobiles to meet their short and fragmented trips, and in turn, weaken their capability to move and inducing restricted activity space.

One of the reasons for the lack of understanding of the determinants and impact of ride-hailing from the gender aspect is the reluctance of TNCs to share their proprietary data (Deka and Fei 2019; Shaheen and Cohen 2019). Thus past studies have to collect individual survey data with small sample sizes, resulting in the difficulty of inferring the entire picture of gender differences in ride-hailing usage as a new mode of urban travel (Habib 2019).

Given the continued uncertainty of gender gap in ride-hailing usage, particularly seeing from a complex bundle of gendered travel patterns and affordability via big data approaches, this study attempts to integrate mobile phone data and a travel survey to answer the following research questions, by selecting Chengdu City, China, as case study city:

(1) Are women dependent on ride-hailing? (Questionnaire).
(2) If women are differentially dependent on ride-hailing as compared to men, how does this gender difference in using ride-hailing occur? (Mobile phone data).

This study contributes to the nexus between gender and transport inequality in the age of platform-based mobility service. it is often assumed that technological innovation in burgeoning smart mobility is the first step toward gender-inclusive transport that will bring equal opportunities to people, facilitate economic efficiency, and promote social development. But only technological advances are insufficient. The aim of this study is to understand women's differentiated travel needs, and thereafter, convert them into urban mobility visions and transport policy making, as equal access to life opportunities for both women and men is only achieved when knowledge of gender disparities has been placed at the center of the governance framework.

10.2 Background

10.2.1 Gendered Travel Needs and Behavior

Gender gap in daily mobility and travel has been an enduring topic at the intersection of transport research, urban studies, and human geography. Scholars have focused on the gender division of labor leading to gender differences in travel demands and temporal-spatial patterns (Kwan 1999; Schwanen et al. 2008). Studies have found that women tend to have shorter travel distance and their daily trips are more fragmented and multi-purpose compared with men (Giuliano 1979; 1983; Scheiner et al. 2011; Scheiner and Holz-Rau 2017). An argument is that a patriarchal society ascribes prime responsibility for domestic labor to women and is likely to force them to seek jobs near their homes (Scheiner and Holz-Rau 2017; Schwanen 2007). This gender division of labor keeps women stuck closer to their homes (Hanson and Johnston 1985; Kwan 1999; Scheiner and Holz-Rau 2017).

Short and fragmented travel needs are difficult to accommodate in public transport designed for connecting work-related hotspots. Women have limited access to flexible travel means (e.g., private cars) to meet their short and fragmentized trip needs. Several reasons are given for women's limited access to private cars. First, gender discrimination in the labor market and women's domestic labor have caused the gender pay gap (Scheiner and Holz-Rau 2017). Lower income causes women's limited capability to afford private automobiles and other expensive but flexible travel means compared with men (Fan 2017). Second, allocation of household car among householders favours men, and thus aggravates a women's access to the car (Scheiner and Holz-Rau 2012; Solá 2016). Third, stigmatization of women as incompetent drivers prevents women from learning to drive and use automobiles (Solá 2016).

The third stream of studies have focused on women's sensitization to safety, resulting in their restricted choice of travel modes, routes and time compared with men (Koskela and Pain 2000; Loukaitou-Sideris 2008; Ceccato and Paz 2017; Zhang et al. 2021). A consensus has been reached on women's vulnerability to offenses, particularly sexual assaults in private/indoor and public/outdoor spheres (Pain 1997; Sweet and Ortiz Escalante 2015). Empirical studies have found that fear of violence has forced women to take precautionary measures to avoid certain 'gendered dangerous' places and mobility means (Loukaitou-Sideris 2014; Ceccato and Paz 2017; Zhang et al. 2021). Thus, fear-induced gender difference in travel behavior is particularly salient after dark (Zhang et al. 2021).

10.2.2 Gender Gap in the Age of Ride-Hailing

An increasing number of studies have observed social equity issues behind the emergence of ride-hailing. Doubt is increasing on the merit of ride-hailing as a transport mode that equally benefits and serves everyone (Ge et al. 2016; Rayle et al. 2016; Deka and Fei 2019). Some studies have used traditional surveys to provide a fundamental understanding of ride-hailing preferences by individuals. Rayle et al. (2016) surveyed 380 ride-hailing users in San Francisco and found that ride-hailing disproportionately served younger residents with higher socioeconomic status. Lavieri and Bhat (2019) found that carpooling mode in ride-hailing is particularly popular with young people (18–44 years) in the Dallas-Fort Worth Metropolitan Area. Alemi et al. (2018) determined that familiarity with modern technology helps younger individuals in California to adopt ride-hailing services at a faster rate than Generation Deka and Fei (2019) found that in the US, young people, people with higher income and education, employed people and people with fewer cars tend to use ride-hailing more often than others.

Apart from age and income, gender has also been examined as a factor for unequal ride-hailing usage. Studies found that women tend to use ride-hailing less than men in the American context (e.g., Deka and Fei 2019). A recent survey showed that women in China are more likely to use ride-hailing than men at night (Qiao and Yeh 2021; Qiao et al. 2023). However, gender has often been taken as a control variable in the literature. Despite a significant yet mixed statistical association, a systematic mechanism of gender difference in ride-hailing accessibility and use remains lacking. Two key questions have been rarely asked: are women dependent on ride-hailing? If ride-hailing serves women differentially, how does this gender difference in using ride-hailing occur? This paper answers these two questions through a preliminary investigation on the interaction of ride-hailing and the two key disadvantages/characteristics of women's daily mobility in ride-hailing: travel demand and affordability.

10.3 Case Study and Data

10.3.1 Case Study City: Chengdu

Chengdu is a sub-provincial city in Southwestern China and serves as the capital of Sichuan province (Fig. 10.1). About 7.74 million people live in the metropolitan areas of Chengdu, which has a built-up area of 802 km^2. Despite a fast-growing metro network (Metro 2020), the demand for reliable and readily available mobility means that overcoming the increased distance to reach jobs and other urban destinations is significant. In this context, platform-based ride-hailing has developed rapidly into a major means of transportation in Chengdu since its first appearance in 2012. DiDi Chuxing (equivalent of Uber) is the largest TNC in China. According to a report from DiDi company, Chengdu has the second highest number of ride-hailing orders amongst all Chinese cities (CBN Data 2016). In the third quarter of 2016, the total number of DiDi smart mobility trips in Chengdu was 130 million, attracting 8.5 million registered users (CBN Data 2016). Chengdu became a model city in terms of promoting legislation on ride-hailing development indicating that the government is favourably inclined towards developing smart mobility (Government 2018). Thus, a favourable governance and policy environment for new modes of transport, as well as citizens' acceptance of new mobility services make Chengdu an excellent case city for studying smart mobility.

This study used a neighborhood (*Shequ*, in Chinese) as the spatial unit of analysis. Neighborhood is the smallest legal self-governing administrative unit of residents in Chinese cities, and is managed by the Community Residents Committee. A neighborhood can be considered as a spatial gathering of the smallest level of residents

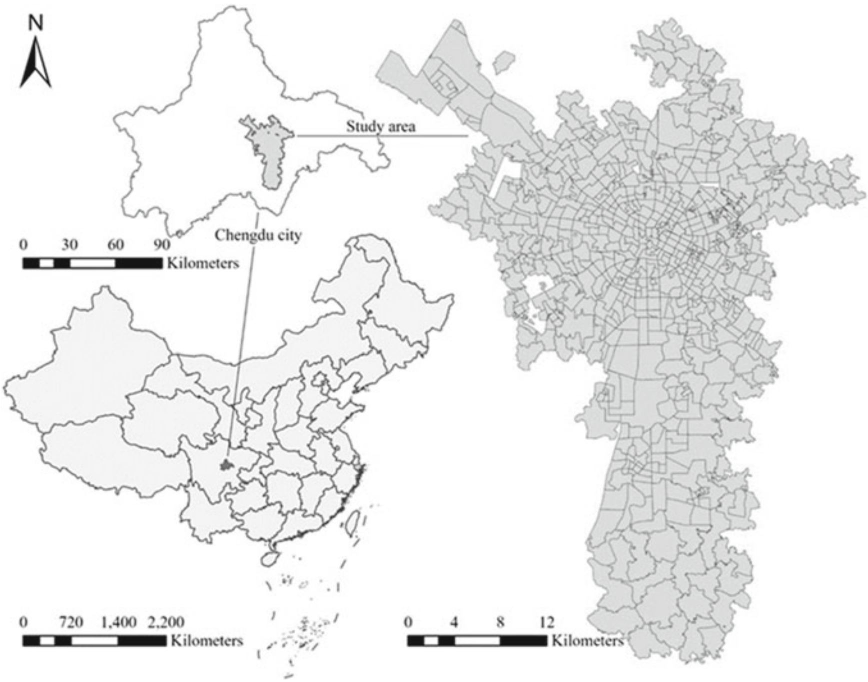


Fig. 10.1 Location of Chengdu and the spatial units of neighborhoods

living nearby. The spatial scale of the neighborhood is similar to the street block in Western cities but typically has a higher population density. The average area of neighborhoods is 1.13 km^2 and the average number of residents is 8,500.

10.3.2 Data Sources

10.3.2.1 Travel Survey

A questionnaire survey was conducted during November 2021 in Chengdu, It was aimed at residents aged 18–60 years who have lived/worked there for over half a year. It did not include students who have highly rigid travel behavior. The survey methods included offline and online questionnaire surveys carried out by adopting a stratified sampling method. For the offline survey, data was collected in the form of intercepted interviews based on the stratified sampling method. For online survey, data were collected in the form of questionnaire links clicked by interviewers in posts placed on online social media communities. A total 1000 of questionnaires were completed by respondents. Eventually, 706 verified questionnaires were included in this study.

Table 10.1 Details of stratified sampling method

Items	Stratified groups	Sampling ratio
Gender ratio	Male: Female	1:1
Regional ratio	Qingyang District: Chenghua District: Jinniu District: Jinjiang District: Wuhou District: Tianfu New District	5:7:6:4:6:4
Age ratio	18–29 years old: 30–39 years old: 40–49 years old: 50–60 years old	6: 4: 4: 3
Offline/Online ratio	Offline: Online	7:3

Table 10.1 summarizes the details of the stratified sampling method according to the seventh national census of Chengdu.

10.3.2.2 Mobile Phone Data

Benefiting from GPS-enabled location tracking on mobile phones, this study also examined factors influencing activity space and affordability, which were hard to obtain from questionnaires. The mobile phone data were obtained from December 1st to 31st, 2018, in Chengdu, from a Chinese mobile company (China Unicom), which is a large operator in Chengdu, occupying at least one-third of the communications market. As a result, 7.74 million residents with their residential coordinates located in the study area were identified with their mobile phone trajectories and usage behavior (e.g., online shopping data, residences, workplaces, and travel time/distance/modes coupled with their gender and age).[1] Residences were assumed to be the place where an individual lived continuously for no less than 14 days at night (20:00–06:00). Major variables of economic status and ride-hailing usage were calculated from the mobile phone trajectories and mobile application usage records. We recorded unique individuals who used ride-hailing applications at least thrice during the recent month as ride-hailing users.[2] 2.21 million residents were thus identified as ride-hailing users. Online consumption was recorded by JD.com, one of the largest e-commerce companies in China. Purchase records from phone users on the website were linked to the mobile phone ID to evaluate the consumption level of owners, so as to represent individual affordability level. To protect personal privacy, this paper did not track individual behaviours. All indicators used in this study were aggregated to the neighborhood level.

[1] Telecom agencies could obtain gender and age from the registered ID card information linking to the mobile phone number.

[2] To identify whether or not mobile phone users are ride-hailing users, we record the communication data between user phone as data sender and ride-hailing companies as data receiver.

10.3.2.3 Geographic Data

Basic geographic data included the point of interests (POI) dataset of built environment (e.g., shops, hospitals, schools, caterings, parks) and transport resources e.g., metro stations and bus stops). This information was provided by Tianditu.com, which is the official online digital map provider in China. A total of 1,037,630 points was contained in the study. The administrative boundary dataset was provided by the land administrative division of China.

10.4 Methodology

10.4.1 Measurement of Activity Space

Activity space represents spatial movement within anurban space and its usage (Nemet and Bailey 2000), and thus, an individual's day-to-day lived experience (Golledge and Stimson 1997). Standard deviational ellipses (SDE) are one of the major approaches to measuring activity space (Yuill 1971). Figure 10.2 illustrates the SDE as an ellipse with major and minor axes that are drawn to represent the magnitude of the minimum and maximum dispersions of a set of points from their mean center. By mapping daily trips and determining the locations of regular activities, SDEs are calculated based on the distance and direction of these locations from home. Activity frequencies can be used to weigh the relative importance of each point (Gesler and Meade 1988).

The following formula is used to measure SDEs in this study:

$$C = \left(\frac{var(x)(x, y)}{(y, x)(y)}\right) = \frac{1}{n}\left(\frac{\sum_{i=1}^{n} \tilde{x}_i^2 \sum_{1}^{n} \tilde{x}_i \tilde{y}_i}{\sum_{i=1}^{n} \tilde{y}_i \sum_{i=1}^{n} \tilde{x}_i^2}\right), \tag{10.1}$$

where

$$(x) = \frac{1}{n}\Sigma_{\mathrm{i}=1}^{n}\left(x_i - \underline{x}\right)^2 = \frac{1}{n}\Sigma_{\mathrm{i}=1}^{n}\tilde{x}_i^2, \tag{10.2}$$

$$(x, y) = \frac{1}{n}\Sigma_{\mathrm{i}=1}^{n}\left(x_i - \underline{x}\right)\left(y_i - \underline{y}\right) = \frac{1}{n}\Sigma_{\mathrm{i}=1}^{n}\tilde{x}_i\tilde{y}_i, \tag{10.3}$$

$$(y) = \frac{1}{n}\Sigma_{\mathrm{i}=1}^{n}\left(y_i - \underline{y}\right)^2 = \frac{1}{n}\Sigma_{\mathrm{i}=1}^{n}\tilde{y}_i^2, \tag{10.4}$$

where x and y are the coordinates of the destination of one personal trajectory i starting from the target neighborhood j, $(\bar{x}, \bar{y})$ represents the mean center of the destinations, and n equals the total number of destinations.

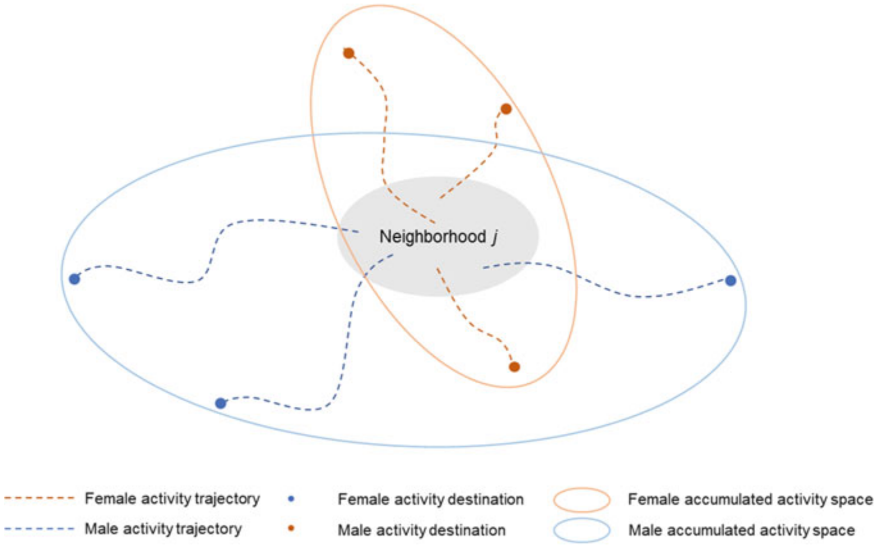


Fig. 10.2 Illustration of standard deviational ellipse-based measurement of accumulated activity space

10.4.2 Built Environment and Socioeconomic Status

Built environment characteristics of neighborhoods have been found to affect the propensity and frequency of ride-hailing use and thus, must be controlled for. Deka and Fei (2019) found a positive association between population and employment densities, and frequency of ride-hailing use. This is consistent with the findings of Hughes and MacKenzie (2016), Wang and Mu (2018) and Alemi et al. (2018). Road density is also found to be positively associated with ride-hailing use (Wang and Noland 2021; Brown 2019). Wang and Noland (2021) found that areas with more retail stores, restaurants and sport and entertainment facilities are associated with higher ride-hailing usage. The findings are consistent with Henao (2017), Rayle et al. (2014) and Wenzel et al. (2019), in that ride-hailing is frequently used for social, leisure and recreational trips. Studies found that neighborhoods with higher automobile ownership and higher income are associated with lower ride-hailing usage (Deka and Fei 2019). Age composition is found to be associated with neighborhood-level ride-hailing usage (Deka and Fei 2019). Young people are more likely to use ride-hailing (Hughes and MacKenzie 2016; Wang and Mu, 2018). Therefore, these built environment and socioeconomic factors are controlled for in our analyses, as presented in Table 10.2.

Table 10.2 Variables in the models

Variables	Variable measures	Data sources
Dependent variable		
Neighborhood-level ride-hailing usage	Number of residents living in this neighborhood that used ride-hailing in one month	25 ride-hailing platform companies
Independent variables		
Female activity space	Accumulated activity space (area of the SDE) of female residents in a neighborhood	China Unicom
Male activity space	Accumulated activity space (area of the SDE) of male residents in a neighborhood	
Female consumption	Accumulated online consumption of female residents in a neighborhood	
Male consumption	Accumulated online consumption of male residents in a neighborhood	
Population density	Residential population in a neighborhood/area of a neighborhood	China Unicom; Gaode Map
Jobs-housing ratio	Number of jobs/number of residents in a neighborhood	
Distance to the city center	Euclidean distance between a neighborhood and the city center (Tianfu Square)	Gaode Map
Distance to metro	Euclidean distance between a neighborhood and the nearest metro station (Tianfu Square)	
Bus stop density	Number of bus stops/area of a neighborhood	
Road density	Length of vehicle roads/area of a neighborhood	
Stores	Density of stores in a neighborhood	
Restaurants	Density of restaurants in a neighborhood	
Sports and entertainment facilities	Density of sports and entertainment facilities in a neighborhood	
Automobile ownership	Number of residents with car ownership/total population in a neighborhood	China Unicom
Ages	Median age of residents in a neighborhood	

10.4.3 Models

Figure 10.3 shows that the distribution of ride-hailing residents in a neighborhood follows the Poisson distribution and has a non-negative characteristic. In this condition, Poisson regression and negative binomial regression (Nbreg) provide efficient coefficients and correct for standard errors (Allison 2012). After an overdispersion test of the dependent variable, the null hypothesis is rejected. Hence, the Nbreg model was used for the final analyses.

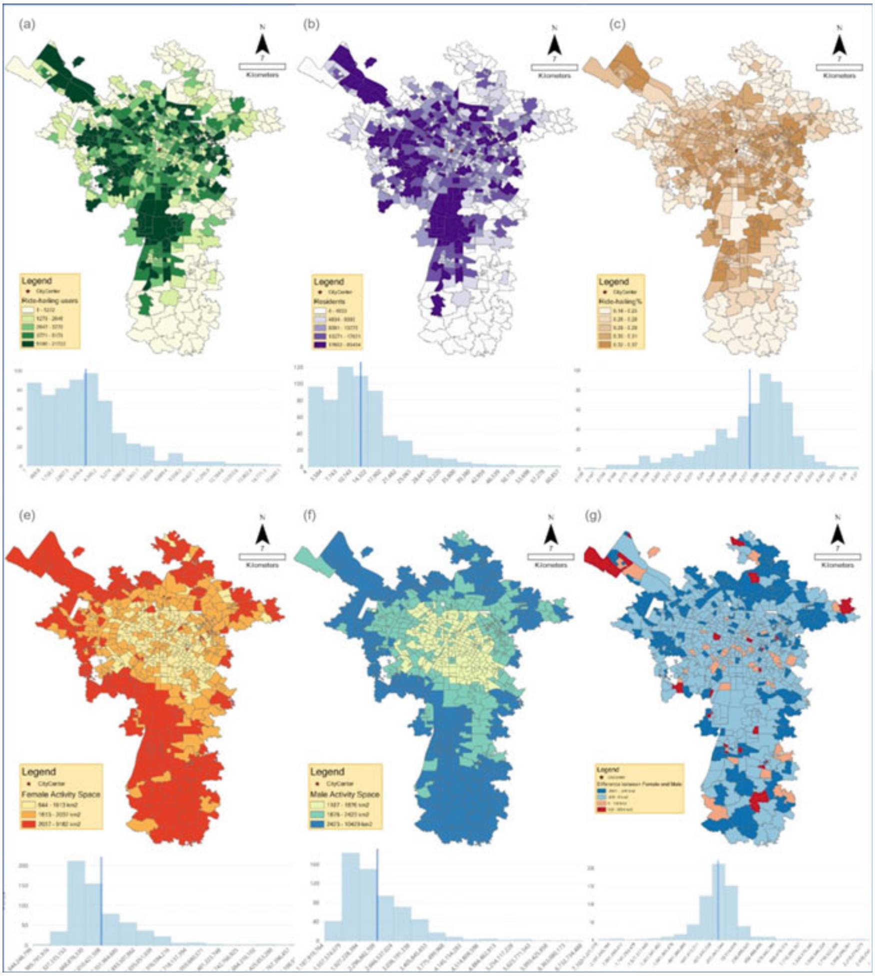


Fig. 10.3 Spatial and statistical distribution of ride-hailing usage and gendered activity space

10.5 Results

10.5.1 Gendered Ride-Hailing Usage Seeing from Travel Survey

Table 10.3 presents the statistics of the travel survey. The survey showed that the majority of the respondents have used ride-hailing at least once for commuting or night entertainment in the past month. No significant gender gap in ride-hailing usage was found. Note that women appear to be more reliant on ride-hailing for nighttime entertainment activities than men. This is consistent with the findings of a study on the spatial distribution of ride-hailing usage in the suburban area (Qiao and Yeh 2021). That is, ride-hailing relatively compensates for women's need to travel safely at night.

The findings on household car allocation in the questionnaire support the conclusion of a mainstream research that women have difficulty obtaining the right to use a household car (Scheiner and Holz-Rau 2012; Solá 2016). When asked if your household had only one private car, would you prefer to allocate it to a male or female member. A total of 656 respondents said it would be allocated to men, while only 50

Table 10.3 Descriptive statistics of travel survey

Variable	Obs	Mean	Std. Dev	Min	Max
Gender (1: Male; 2: Female)	706	1.53	0.50	1.00	2.00
Commuting by RH	680	0.26	0.44	0.00	1.00
Male	329				
Female	351				
Night entertainment by RH	578	0.66	0.47	0.00	1.00
Male	271				
Female	307				
Number of household car:	706	0.68	0.61	0.00	3.00
None	277				
Only one	379				
Two	48				
Three	2				
Allocate car to:	706	1.07	0.26	1.00	2.00
Male member	656				
Female member	50				
Personal income	706	6786.12	4698.98	0	50,000
Male income	329	7462.01	5201.36	3000	50,000
Female income	377	6196.29	4129.38	0	50,000
Household income	706	21,974.50	13,160.52	5000	>50,000

respondents said it would be allocated to women. This result suggests that women in our case study city have less access to cars.

Women also reported lower personal income than men. The average income of women is about 83% of that of men. These findings are in line with previous research in China, that women are paid 75.4% of what men are paid, though the gap in labor force participation rate between women (60.5%) and men (75.0%) is not that sharp (Xiu and Gunderson 2013). This gap exists across age groups, educational levels and regions (Bai et al. 2022). Average monthly income in Chinese mega-cities is around 6,000 CNY (CSB 2020), while the cost of an average ride-hailing trip is between 22 and 43 CNY (CBN Data 2016) Thus, a ride hailing trip consumes a greater share of a woman's personal income. This difference in income may hinder women's ride-hailing usage.

10.5.2 Spatial Variations in Ride-Hailing Usage Seeing from Big Data

Figure 10.3a illustrates the statistical distribution of, and spatial variations in, the number of residents using ride-hailing at a neighborhood scale. In general, the closer the neighbourhood is to the central business districts, the lower the number of residents using ride-hailing in that neighborhood. Figure 10.3b shows the spatial and statistical distribution of the number of residents in a neighborhood. Figure 10.3c shows the spatial and statistical distribution of ratio of ride-hailing users to residents in a neighborhood.

10.5.3 Gendered Activity Space

Figure 10.3d and e illustrate the average activity space of female and male residents, respectively, in 614 neighborhoods. On average, the activity space for females (1.993 km^2) is smaller than that of the males' (2.311 km^2). The direct Euclidean distance between the workplace and residence for women is 9.7 km, whereas it is 10.4 km for men. In general, the closer a neighbourhood is to the central business district, the lower the average activity space of its residents.

Figure 10.3g illustrates the gender gap in the activity space of 614 neighborhoods in Chengdu. The average activity space of male residents was larger than female residents in 89% of neighborhoods. In general, inner-city neighborhoods tended to have a smaller men-to-women gap in average activity space than suburban neighborhoods.

These numbers are a reflection of the gender division of labor, in which women tend to engage in domestic labor in and around their neighborhoods, whereas men are willing to work father away from their neighborhoods for opportunities (Kwan 1999; Schwanen et al. 2008; Shen et al. 2021).

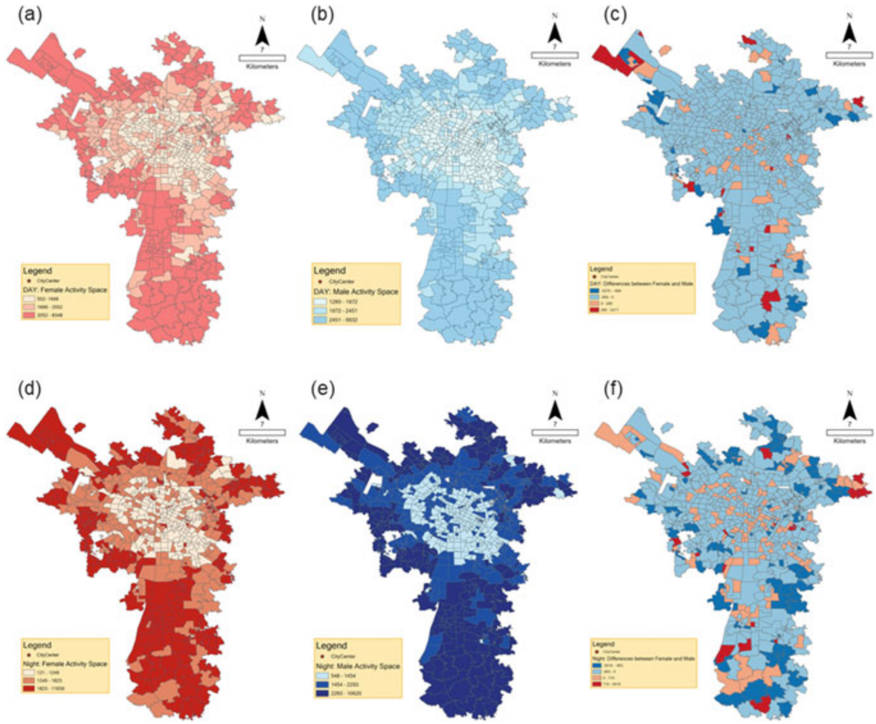


Fig. 10.4 Gendered activity space in neighborhoods by time

Figure 10.4a–c show the activity spatial distribution of daytime travel (8 a.m. to 8 p.m.). Figure 10.4d–f show the activity spatial distribution of night travel (8 p.m. to 8 a.m.). Night travel produced less difference in the activity space between women and men than their daytime travel. Coupled with the results of the questionnaire survey, the narrowing of the night-time activity space gap may benefit from women's access to ride-hailing at night.

10.5.4 Ride-Hailing Usage, Gender, and Influencing Factors

The regression result of model 1 is presented in Table 10.4, which shows that accumulation of female residents' *activity space* is positively associated with the number of residents using ride-hailing in a neighborhood **(29.366**)**. Meanwhile, the accumulation of male residents' *activity space* is not significantly associated with the number of residents using ride-hailing in a neighborhood. These results indicate that female's daily activity space significantly determined the use of ride-hailing while males' does not.

Table 10.4 Result of model 1

Residents used ride-hailing services	Coef.	Std. err.	irr.	Sig.	VIF
Female activity space (log)	29.366	14.494	0.567	**	4.15
Female activity space ^2	−0.663	0.339	0.515	*	/
Male activity space (log)	−0.668	17.266	0.513		4.78
Male activity space ^2	0.029	0.400	1.029		/
Population density	0.000	0.000	1.000	***	1.97
Jobs-housing ratio	−0.931	0.147	0.394	***	1.13
Distance to city center	0.000	0.000	1.000	***	3.39
Distance to metro	0.000	0.000	1.000	***	1.49
Bus stop density	0.012	0.004	1.012	***	1.50
Shopping mall and restaurant density	0.000	0.000	1.000		2.77
Living services density	−0.001	0.000	0.999	***	4.39
Sport and entertainment facilities density	0.001	0.001	1.001		2.40
Automobile ownership ratio	−9.517	1.95	0.000	***	1.52
Age (median)	0.025	0.009	1.026	***	1.18
Constants	−314.996	107.14	0.000	***	
lnalpha	−0.623	0.053	−0.623		
Number of *obs*	614	Akaike crit. (AIC)		11,031.030	
Chi-square	264.679	Bayesian crit. (BIC)		11,101.749	
Prob > chi2	0.000				

*** p<.01, ** p<.05, * p<.1

Table 10.5 presents that the regression results of model 2, which shows that female residents' accumulated consumption is positively associated with neighborhood-level ride-hailing usage **(0.681**)**, while male residents' accumulated consumption is not. Hence, females' income is an influencing factor that could prevent their access to ride-hailing usage while males' does not.

10.6 Conclusions and Discussion

The burgeoning of platform-based ride-hailing has given rise to numerous debates in the recent decade. The view that ride-hailing will bring equal opportunities to people, facilitate economic and social efficiency, and promote social development gained popularity among technology companies, investors, and transport agencies (Cohen and Shaheen 2018). However, an increasing number of scholars have challenged this view by unpacking the exclusive and unequal nature of everyday ride-hailing use (Ge et al. 2016; Deka and Fei 2019). These studies called for attention to the unequal

Table 10.5 Result of model 2

Residents used ride-hailing services	Coef.	Std. err.	irr	Sig.	VIF
Female consumption	0.681	0.317	1.975	**	3.84
Female consumption^2	−0.204	0.075	0.815	***	/
Male consumption	0.437	0.318	1.548		4.04
Male consumption^2	−0.099	0.076	0.906		/
Population density	0.000	0.000	1.000	***	2.26
Jobs-housing ratio	−0.692	0.159	0.500	***	1.18
Distance to city center	0.000	0.000	1.000		2.40
Distance to metro	0.000	0.000	1.000	***	1.51
Bus stop density	0.012	0.004	1.012	***	1.50
Shopping mall and restaurant density	0.000	0.000	1.000		2.77
Living services density	−0.001	0.000	0.999	***	4.43
Sport and entertainment facilities density	0.001	0.001	1.001		2.40
Automobile ownership ratio	−10.396	1.904	0.000	***	1.60
Age (median)	0.030	0.010	1.031	***	1.18
Constant	7.332	0.623	1,529.067	***	
lnalpha	−0.558	0.053	−0.558		
Number of *obs*	614	Akaike crit. (AIC)		11,077.840	
Chi-square	217.869	Bayesian crit. (BIC)		11,148.560	
Prob > chi2	0.000				

*** p<0.01, ** p<0.05, * p<0.1

benefits and externalities of ride-hailing among social groups, and how the rapid development of ride-hailing is involved in the exacerbation of transport inequality. This study joins this debate by presenting the gendered nature of everyday mobility and their ride-hailing usage. Two key questions are answered: Are women dependent on ride-hailing? If ride-hailing serves women differentially, how does this gender difference in ride-hailing use occur?

We applied an innovative combination of big data and travel survey to answer the preceding questions by taking Chengdu, China, as a case study. The survey results and modelling analysis suggest following conclusions. (1) On average, women have lower personal income and less access to private cars than men. (2) Women are more likely to use ride-hailing than men for their night travel. (3) Given that women lack travel alternatives (e.g., private car) for trips beyond their neighborhoods and have lower income and less access to flexible travel means (e.g., driving), ride-hailing could mitigate the gender gap in daily travel (4) Lastly, gendered travel needs (activity space) and gendered affordability are two major factors influencing women's ride-hailing usage.

We argue that the emergence of ride-hailing improves a women's ability to move and mitigates gender gap in transport at an aggregate level in Chinese megacities. However, the nature of gender gap in everyday travel is caused by low income and less access to cars, which are the results of women being disadvantaged in the division of labor. This long persisting and structural barrier in society cannot be solved immediately. Results indicate that ride-hailing usage increases with women's activity space and affordability, ride-hailing may benefit middle-income and employed women (enhanced their accessibility to jobs and social opportunities) more than a whole group of women across a spectrum of socio-economic status. Low-income housewives remain excluded from this technological advancement. Thus, TNCs should be encouraged to diversify service provisions according to the travel demands of different social groups. A possible solution is a special subsidy scheme motivating drivers to provide services to short-distance passengers within and around neighborhoods to provide lower starting fares.

Further, some limitations of this research are acknowledged. This study takes women as a whole to develop a comparison between men and women. This women–men dichotomy is challenged by intersectionality theory, which calls for attention to internal differences in gendered experiences within women (Valentine 2007). Differences in agency and capability in determining emerging social opportunities and mitigating gender disadvantage among women as shaped by class, age and other socio-economic factors are emphasized (Collins 2000; McCall 2009). Socio-institutional milieu matters to women's experiences with respect to patriarchy, and should also be examined. Ride-hailing may serve men more than women, thereby enlarging gender gap in regions with low labor participation rate and lower income levels for women, or even affected by the culture of society. Therefore, a context-specific analysis is needed to understand how ride-hailing affects gender inequality in daily travel.

Acknowledgements This study received institutional ethics approval from the University of Hong Kong, Reference HREC: EA210312). This work is an extension from "Mind the gender gap in ride-hailing from demand side" (Qiao, Zhang, and Yeh 2023). We would like to thank the financial support from the Dissertation Fellowship of Peking University-Lincoln Institute Center for Urban Development and Land Policy (Grant Number DS04-20211001-QS); Chan To-Haan Endowed Professorship Fund of the University of Hong Kong; Joint Programming Initiative (JPI) Urban Europe and National Natural Foundation of China (NSFC) (Grant Number: 71961137003); and Guangdong–Hong Kong-Macau Joint Laboratory Program of the 2020 Guangdong New Innovative Strategic Research Fund, Guangdong Science and Technology Department (Project No.: 2020B1212030009).

References

Alemi F, Circella G, Handy S, Mokhtarian P (2018) What influences travelers to use Uber? exploring the factors affecting the adoption of on-demand ride services in California. Travel Behav Soc 13:88–104

Bai W, Yan-Li Lee YL, Liao J, Wu L, Xie M, Zhou T (2022) The Gender Pay Gap in China: insights from a discrimination perspective. arXiv:2206.09306

Brown A (2019). Redefining car access: ride-hail travel and use in Los Angeles. J Am Plann Assoc 85(2):83–95

Ceccato V, Paz Y (2017) Crime in São Paulo's metro system: sexual crimes against women. Crime Prevent Commun Saf 19(3):211–226

China's Statistical Bureau (2020) China Statistical Yearbook 2019, China Statistical Publishing House, Beijing (2020)

Clewlow RR, Mishra GS (2017) Disruptive transportation: the adoption, utilization, and impacts of ride-hailing in the United States

Cohen A, Shaheen S (2018) Planning for shared mobility. American Planning Association, Chicago IL, USA

Collins PH (2000) Black feminist thought: knowledge, consciousness, and the politics of empowerment, 2nd edn. Routledge, NY

Contreras SD, Paz A (2018) The effects of ride-hailing companies on the taxicab industry in Las Vegas, Nevada. Transp Res Part A Policy pract, 115, 63–70

CBN Data (2016) Industry observation of shared mobility. https://cbndata.com/report/112?isReading=report&page=4, Accessed 12 Oct 2021

Deka D, Fei D (2019) A comparison of the personal and neighborhood characteristics associated with ridesourcing, transit use, and driving with NHTS data. J Transp Geogr 76:24–33

Fan Y (2017) Household structure and gender differences in travel time: spouse/partner presence, parenthood, and breadwinner status. Transportation 44(2):271–291

Ge Y, Knittel CR, MacKenzie D, Zoepf S (2016) Racial and gender discrimination in transportation network companies(No. w22776). National Bureau of Economic Research

Gesler WM, Meade MS (1988) Locational and Population Factors in Health Care-seeking Behavior in Savannah, Georgia. Health Serv Res 23(3):443–462

Giuliano G (1979) Public transportation and the travel needs of women. Traffic Q 33:607–616

Giuliano G (1983) Getting there: women and transportation. the technological woman: interfacing with tomorrow. Praeger, New York

Golledge RG, Stimson RJ (1997). Spatial behavior: a geographic perspective. Guilford, NewYork

Habib KN (2019) Mode choice modelling for hailable rides: an investigation of the competition of Uber with other modes by using an integrated non-compensatory choice model with probabilistic choice set formation. Transp Res Part A Policy Pract 129, 205–216

Hanson S, Johnston I (1985) Gender differences in work-trip length: explanations and implications. Urban Geogr 6:193–219

Henao A (2017) Impacts of ridesourcing-lyft and uber-on transportation including VMT, mode replacement, parking, and travel behavior. PhD thesis, University of Colorado, Boulder.

Hughes R, MacKenzie D (2016) Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators. J Transp Geogr 56:36–44

Jin ST, Kong H, Wu R, Sui DZ (2018) Ridesourcing, the sharing economy, and the future of cities. Cities, 76, 96–104

Kong H, Zhang X, Zhao J (2020) How does ridesourcing substitute for public transit? a geospatial perspective in Chengdu China. J Transp Geogr 86:102769

Koskela H, Pain R (2000) Revisiting fear and place: women's fear of attack and the built environment. Geoforum 31(2):269–280

Kwan M-P (1999) Gender, the home-work link and space–time patterns of nonemployment activities. Econ Geogr 76(4):370–394

Lavieri PS, Bhat CR (2019) Modeling individuals' willingness to share trips with strangers in an autonomous vehicle future. Transp Res Part A Policy pract, 124, 242–261

Loukaitou-Sideris A (2014) Fear and safety in transit environments from the women's perspective. Secur J 27(2):242–256

Loukaitou-Sideris AL (2008) How safe is the ride? Evaluation of design and policy responses to women's fear of victimization and crime

McCall L (2005) The complexity of intersectionality. Signs 30:1771–1800

Nemet GF, Bailey AJ (2000) Distance and health care utilization among the rural elderly. Soc Sci Med 50:1197–1208

Qiao S, Yeh AGO (2021) Is ride-hailing a valuable means of transport in newly developed areas under TOD-oriented urbanization in China? evidence from Chengdu City. J Transp Geogr 96:103183

Qiao S, Yeh AGO, Zhang M (2020) Effects of State-led Suburbanization on traffic crash density in China: evidence from the city proper of Chengdu. Accid Anal Prev 148(105775)

Qiao S, Zhang M, Yeh AGO (2023) Mind the gender gap in ride-hailing from the demand side. J Transp Geogr 107:103531

Rayle L, Shaheen S, Chan N, Dai D, Cervero R (2014) App-based, on-demand ride services: Comparing taxi and ridesourcing trips and user characteristics in san francisco university of california transportation center (uctc). University of California, Berkeley, United States 2:49–52

Rayle L, Dai D, Chan N, Cervero R, Shaheen S (2016) Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transp Policy 45:168–178

Scheiner J, Holz-Rau C (2012) Gendered travel mode choice: a focus on car deficient households. J Transp Geogr 24(2):250–261

Scheiner J, Holz-Rau C (2017) Women's complex daily lives: a gendered look at trip chaining and activity pattern entropy in Germany. Transportation 44(1):117–138

Scheiner J, Sicks K, Holz-Rau C (2011) Gendered activity spaces: trends over three decades in Germany. Erdkunde 65(4):371–387

Schwanen T (2007) Gender differences in chauffeuring children among dual-earner families. Profess Geogr 59(4):447–462

Schwanen T, Kwan MP, Ren F (2008) How fixed is fixed? gendered rigidity of space–time constraints and geographies of everyday activities. Geoforum 39(6):2109–2121

Shaheen S, Chan N, Bansal A, Cohen A (2015) Shared mobility: a sustainability and technologies workshop: definitions, industry developments, and early understanding. UC Berkeley: Transp Sustain Res Center. Retrieved from https://escholarship.org/uc/item/2f61q30s

Shaheen S, Chan N (2016). Mobility and the sharing economy: potential to facilitate the first-and last-mile public transit connections. Built Enviro 42(4): 573–588

Shaheen S (2018) Shared mobility: the potential of ridehailing and pooling, Three Revol 55–76. Island Press, Washington, DC

Shaheen S, Cohen A (2019) Shared ride services in North America: definitions, impacts, and the future of pooling. Transp rev 39(4): 427–442

Shen Y, Ta N, Liu Z (2021) Job-housing distance, neighborhood environment, and mental health in suburban Shanghai: a gender difference perspective. Cities 115:103214

Solá AG (2016) Constructing work travel inequalities: the role of household gender contracts. J Transp Geogr 53, 32–40

Sweet EL, Ortiz Escalante S (2015) Bringing bodies into planning: visceral methods, fear and gender violence. Urban Stud 52(10):1826–1845

Valentine G (2007) Theorizing and researching intersectionality: a challenge for feminist geography. Prof Geogr 59(1):10–21

Wang M, Mu L (2018) Spatial disparities of Uber accessibility: an exploratory analysis in Atlanta, USA. Comput Environ Urban Syst 67:169–175

Wang S, Noland RB (2021) Variation in ride-hailing trips in Chengdu, China. Transp Res Part D Transp Environ 90:102596

Ward JW, Michalek JJ, Azevedo IL, Samaras C, Ferreira P (2019) Effects of on-demand ridesourcing on vehicle ownership, fuel consumption, vehicle miles traveled, and emissions per capita in US States. Transp Res part C Emerg technol, 108, 289–301

Wenzel T, Rames C, Kontou E, Henao A (2019) Travel and energy implications of ridesourcing service in Austin, Texas. Transp Res Part D Transp Environ 70:18–34

Xiu L, Gunderson M (2013) Gender earnings differences in China: base pay, performance pay, and total pay. Contemp Econ Policy 31(1):235–254

Young M, Farber S (2019) The who, why, and when of Uber and other ride-hailing trips: an examination of a large sample household travel survey. Transp Res Part A Policy and pract, 119, 383–392

Yuill RS (1971) The standard deviational ellipse: an updated tool for spatial description. Geografiska Annaler Ser b Human Geogr 53B:28–39

Zhang M, Zhao P, Qiao S (2020) Smartness-induced transport inequality: privacy concern, lacking knowledge of smartphone use and unequal access to transport information. Transp Policy 99:175–185

Zhang M, Zhao P, Tong X (2021) Constructing women's immobility: fear of violence and women's constricted nocturnal travel behaviour. Travel Behav Soc 26:178–192

Zhang M, Zhao P (2021) Literature review on urban transport equity in transitional China: from empirical studies to universal knowledge. J Transp Geogr 96(20121):103177

Chapter 11
Urban Airspace Route Planning for Advanced Air Mobility Operations

Xi Wang, Perry Pei-Ju Yang, Michael Balchanos, and Dimitri Mavris

Abstract This chapter aims to explore methods and procedures for route planning in urban air space for Advanced Air Mobility (AAM) operations using a 3D GIS environment. Route planning for urban air space through data analytics is produced to support planners in decision-making by visualizing the key influential factors in a 3D urban environment. Having defined two use cases for the City of Atlanta, different data types are needed to account for those factors and are introduced and tested. The use case for the Atlanta Aerotropolis has represented the need to plan a short-distance inner-city AAM network to serve as a public transportation network. In contrast, the other use case is a remote terminal shuttle service for the Atlanta airport. Recommendations on validating and optimizing the AAM networks are made at the end of this chapter.

Keywords Advanced air mobility · 3D GIS · Electric vertical take-off and landing aircraft

11.1 Introduction and Motivation

An Advanced Air Mobility (AAM) system is a new type of transportation method that emerged in the past decade. It is a concept developed on top of the traditional Urban Air Mobility (UAM), an air mobility system within the urban limit with usually less than 100 miles of service distance. Instead of focusing on passenger

X. Wang (✉) · P. P.-J. Yang · M. Balchanos · D. Mavris
Georgia Institute of Technology, 225 North Avenue, Atlanta, GA 30332, USA
e-mail: xwang3106@gatech.edu

P. P.-J. Yang
e-mail: perry.yang@design.gatech.edu

M. Balchanos
e-mail: michael.balchanos@asdl.gatech.edu

D. Mavris
e-mail: dimitri.mavris@aerospace.gatech.edu

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_11

and cargo transfer within a city, the AAM concept incorporates long-range flights up to traditional inter-city regional flights, public service flights such as ambulance or police flights, and private recreational flights (Federal Aviation Administration 2022a).

Like UAM, the AAM system utilizes electric Vertical Take-off and Landing (eVTOL) aircraft as its primary technology enabler and air vehicle configuration. By using these helicopter-like air vehicles, AAM goes beyond traditional road-based transportation planning that is constrained by operating on a 2D plane. Redirecting a portion of the traffic into the air can reduce the load in congested road networks in urban areas. eVTOL aircraft use also reduces carbon emissions as they are powered by battery or hydrogen fuel cells and are responsible for zero direct emissions to the atmosphere during operation. During the past few years, this industry has received attention as operating eVTOLs are being produced and undergoing test flights.

The metropolitan areas in the States face road network congestion, and their residents spend a substantial amount of time and money on the road in traffic jams. Atlanta ranked high on the most congested cities in the United States (INRIX 2021). According to INRIX, an average Atlanta resident would spend 53 h a year being stuck in a traffic jam and would therefore waste \$800 due to the time lost and extra fuel burnt. Additionally, with heavy congestion, alternative transportation methods that utilize the road network (such as ride-share or metro buses) will also be impacted, rendering them ineffective. When a vehicle with a traditional Internal Combustion Engine (ICE) stops in congested traffic, idles the engine, then accelerates to move again, more greenhouse gases will be emitted from the tail pipe compared to when the roads are not congested (Vehicle Technology Office 2015). 30 million tons of CO_2 are emitted into the atmosphere due to the unnecessary idling of road vehicles every year in the United States. While the congestion is worsening, cities are imposing emission goals to reduce greenhouse gas emissions. Atlanta's Transportation Plan stated that the city plans to achieve the goal of reducing emissions from transportation-related activities by 40% compared to 2009 levels by 2030. The plan stated that single occupancy vehicle travel is to be reduced as well (American Society for an Energy Efficient Economy 2020).

A well-planned and implemented AAM system could help with the situation by reducing both traffic congestion and emissions. With congested traffic on the ground emitting greenhouse gases, the AAM offers a new perspective to solving the problem as eVTOL aircraft are available to transport passengers from city areas to their destination in the air. The system operates from vertiports designed especially for eVTOL aircraft. Similar to a helicopter landing pad, these vertiports that can be built on top of office buildings or city parking decks enable the ability for the AAM network to reach into the most congested city centers where the demand for travel is the highest. However, unlike traditional helicopters, which are environmentally unfriendly due to the high level of fuel consumption, and unfeasible for most travelers due to the expensive operating cost, eVTOL aircraft are free of direct emission and are able to operate at a much more feasible price for commuters.

With the high flexibility for placing the vertiports and the technology advantage of using eVTOL aircraft, the presence of an AAM system has great potential to reduce the number of privately owned vehicles and vehicles-for-hire used by daily commuters and airport travelers. With the new transportation system, a portion of the population may choose to commute and travel to the airport by eVTOL aircraft.

As the number of ICE-powered vehicles decreases, the emission of greenhouse gases and carbon dioxide related to airport traffic also decreases.

An airport could greatly benefit from an AAM system as the use of eVTOL aircraft would allow the use of remote terminals. Delta Airlines had recently invested 60 million dollars in Joby Aviation to expedite travel times to the airport (Bogaisky 2022). Passengers would be able to check in and drop off their luggage at vertiports in the city centers or other points of interest far away from the airport prior to the traditional check-in time and enjoy the traffic-free flight on an eVTOL aircraft to the airport. Passengers will also be able to park their vehicles in the parking lots next to the vertiport instead of airport parking decks.

An airport is a central hub of both ground and air traffic. Using the airport as a use case to develop the route planning approach will consider more factors and be robust to apply in other less complex urban contexts. Therefore, an inter-city AAM network as an airport shuttle for the Hartsfield-Jackson Atlanta International Airport is selected to be a case study for this method.

Atlanta has an Aerotropolis planned around the Hartsfield Jackson Atlanta International Airport. Many cities are turning their city airports into airport cities, or aerotropolis, in which city centers are built around globally significant airports (Kasarda 2013). It is an alignment of the metropolitan region to better leverage an airport's assets and provides a framework for the strategic planning and development of economic activity and real estate (Atlanta Aerotropolis 2016). It includes the airport, its surrounding areas, businesses, residential areas, and all necessary facilities required for daily activity. An inner-city AAM network as public transportation within the aerotropolis catering to the daily commuter is selected to be a use case.

Despite all the potential benefits an AAM system could bring to the public transportation network for the local residents, there is uncertainty on how to plan the routes for such a system on a city-wide scale. There exists a gap between the perspectives of existing studies from aerospace engineers and the perspective required by city planners, which this research tries to bridge. Due to the many existing types of air traffic, implementing an AAM system in an airport's complex multi-dimensional traffic network would be difficult. A planning tool for 3D urban airspace utilizing computing technology in a digital environment will enable city planners to address public transportation networks both on the ground and in the air. As the concept of eVTOL and AAM are relatively new, there are limited studies on route planning for such transportation systems. Current existing studies for AAM systems use simplified ways to site vertiports.

11.1.1 Literature Review

Past studies and simulations for AAM systems and eVTOLs usually planned routes in a less sophisticated manner. Larger scale simulations avoided detail planning and used straight lines connecting the center of Census Tract (CT) blocks as routes for the simulated AAM networks (Chitale et al. 2021). Even in studies with sophisticated routing, the vertiport siting is still highly simplified. A famous AAM demand study from the Virginia Institute of Technology located all vertiports in the center of CT blocks for simplification (Rimjha et al. 2021). The result from this study is still actively being used by NASA for ongoing research on the AAM system feasibility.

One study from MIT looked into planning the shortest route between vertiports separated by controlled airspace using computer algorithms (Vascik et al. 2020). The study sug- gested an approach to account for different airspace and Air Traffic Control (ATC) operations using a case study of the San Francisco Bay Area. This study assumes that the AAM network operates within the existing airspace structure and follows existing regulations. Therefore, the result is a rapidly changing routing AAM net- work depending on the position of all other aircraft in the airspace. However, with the introduction of the Federal Aviation Administration (FAA) 's UAM Concept of Operations (CONOPS) in late 2020, AAM operations are to be conducted within eVTOL corridors established beforehand (Federal Aviation Administration 2020). The research for urban airspace route planning for AAM operations is urgently needed.

This chapter provides a method that focuses on planning the route based on carefully selecting the vertiport site for an AAM system. Instead of the traditional method of using the geographic center of CT blocks as vertiport locations, a more comprehensive and realistic method to site them is provided. This is done by study- ing the potential factors contributing to the vertiport siting and the route planning decision-making process. A GIS environment visualizing those factors was created to support the planners in finding the best solution. Factors affecting vertiport selec- tion are crucial as the location determines the demand for the AAM service, how the vertiport affects the residents, and so on. The factors affecting the route are focused on the airspace regulations and laws by the FAA (Federal Aviation Administration 2020; Designation of Class A B, C, D, E 2022; Special Use Airspace 2022; Airport Noise Compatibility Planning 2022).

11.2 AAM and eVTOLs

The AAM system has received a surge in attention during the past few years as eVTOLs are being produced and tested all around the globe by aerospace companies. However, there are still many gaps to be bridged before the AAM system can be practically useful. American Institute of Aeronautics and Astronautics (AIAA) formed an AAM task force to push the development of AAM systems, and the task force identified several gaps to be addressed in both the near and far term. These gaps include critical issues such as the need for low-noise vehicles, lack of vertiport safety designs, uncertainty during integration with other travel methods for ultimate efficiency, and shortage of pilots and technical labor as the AAM industry uses high training standards from the traditional aviation industry, etc. These are a fraction of the many technical gaps that engineers and designers have to face before the AAM system could be put into practical use.

Regulators such as the FAA are building new regulations to accommodate the new system. As safety is their top priority, the speed of regulatory approval for the AAM system will be slow. With the high uncertainty on eVTOL technology, regulators have had regulations that slow down the development. Regulations for pilot training requirements are an example. Even with the advance in autonomy in eVTOL aircraft, pilots are subject to standards of training much higher than traditional airline pilots in multiple types of aircraft (fixed-wing and helicopter) before they are able to operate a highly autonomous eVTOL aircraft. This is due to the fact that eVTOL aircraft are categorized as powered-lift by regulators, a mid-point between a helicopter and a fixed-wing aircraft.

The gaps are not limited to the technical and regulation side. Even if the system is perfectly designed, the public will need time to build trust to be willing to ride in the AAM system. If a fatal crash happened in the early phase of commercialization, it could be a major setback for public acceptance. The announcement of using the AAM system as a transportation method at major global events such as the Paris Olympic 2024 and the Los Angeles Olympic 2028 to bring the system into public view is surely a first step towards broader public acceptance.

11.2.1 eVTOL Aircraft Capabilities

This chapter uses the standard eVTOL aircraft configurations to define them in three categories. These configurations dramatically affect their average range, speed, and endurance capabilities.

A tilt-rotor design aims to operate for the longest distance between landings (typical range 250 km). It is the fastest configuration at around 300 km/h. A tilt- rotor eVTOL rotates its rotors upward to point the thrust downward during the take-off and landing phases, redirect the thrust forward, and generate lift using its conventional wings during the cruise phase. Less power is needed during the cruise phase as the aircraft operates as a traditional fixed wing, and the rotors only generate forward thrust. It is also the fastest configuration due to no unused rotor generating aerodynamic drag. However, the complex tilting mechanism is expensive and is prone to mechanical failure.

A lift and cruise design use two sets of rotors, one set to generate lift during take-off and landing and another set to provide forward thrust during cruise. A conventional wing is used to produce lift during cruise. This design does not require the complex mechanical design for the tilting mechanism but carries dead weight and

Fig. 11.1 Conceptual eVTOL designs from NASA

aerodynamic drag during the cruise stage as the lifting rotors are not being used. This configuration typically is slower and has a shorter range than the tile-rotor eVTOLs. A multi-rotor design uses multiple upward-pointing rotors to generate lift and augment the individual thrust on each motor to vector the overall thrust. With a look similar to a multi-rotor drone, this configuration is the simplest in design. There is no conventional wing to generate lift during cruise. Therefore, the rotors operate at high power throughout the entire flight. It is also the easiest to pilot as there is no transition phase from cruise to hover, giving it the highest agility. This configuration is the most reliable as it does not have tilting mechanisms but is the slowest and has the shortest range.

Figure 11.1 shows three deference conceptual designs by NASA. From left to right shows the tilt-rotor configuration, lift and cruise configuration, and multi-rotor configuration.

Depending on the size of the AAM network, different configurations should be considered. Long-range regional inter-city commute networks similar to the current helicopter operation should consider the tilt-rotor category due to the faster speed, and short-range urban inner-city operations such as air taxis could use multi-rotor eVTOLs due to the high agility and less complex mechanical systems. For the Atlanta study case, a short-range AAM system is proposed and so multi-rotor eVTOLs would most likely be used. Therefore, the shorter range for multi-rotor eVTOLs (less than 70km) would be a constraining factor when deciding the spacing of vertiports as a profitable AAM network that would require eVTOLs to complete multiple legs before being stored for charging.

11.2.2 Concept Overview

As shown in Fig. 11.2, a high-level operational concept diagram defines the route planning problem. All factors shown in the diagram are needed to plan an AAM network and must be considered for decision-making.

Before planning any route, the locations of the vertiports in the AAM network are selected. The chapter looks into multiple areas where there exist factors that will impact the practicality of the location for a vertiport to be built on. The chapter looks first into the factors regarding air operations, from eVTOL aircraft capabilities to the

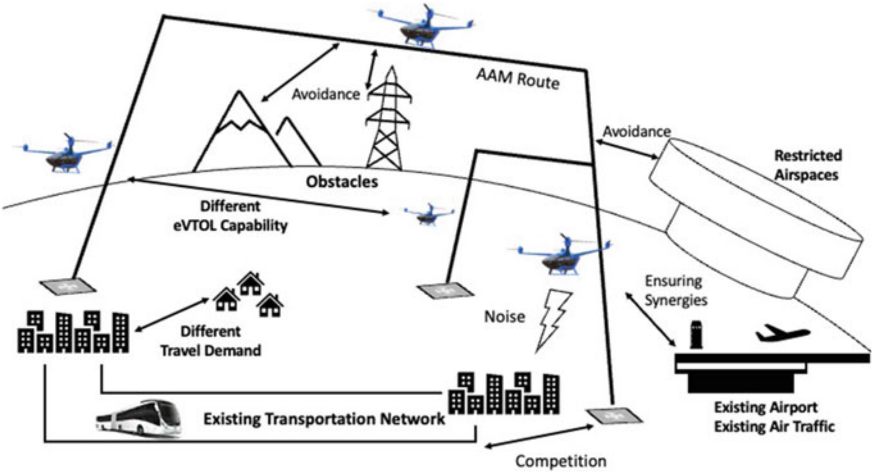


Fig. 11.2 High-level operational concept for AAM route planning

FAA regulation on airspace usage, then the other factors, such as passenger travel demand and potential competition from existing public transportation.

After the locations of the vertiports are selected, the routes to connect them are the next to be determined. A straight line would be the most ideal as it minimizes the time required to reach destinations. However, it interferes with the factors discussed in the following section, and modifications must be done.

11.3 Use Case: Atlanta Airport City

In the case study, routes were proposed for the two networks in the Atlanta use case. First, the Inner Aerotropolis Network is considered based on factors including travel demand, existing public transportation, regulations, obstacles, and environmental impacts such as noise for the urban air transportation system. Second, the Airport Suttle Network connects the airport to areas of high travel demand. Using the concept of remote terminals, the vertiports in this network could reduce the process of passenger security screening. All factors affecting decision-making were visualized in the GIS model.

11.3.1 Inner Aerotropolis Network

The first use case provides a route planning analysis and decision for the Inner Aerotropolis Network.

11.3.1.1 Travel Demand

The most important factor for the entire system is the travel demand for such an urban air transportation system. People flow, household income, and competition from existing traffic networks are considered in the chapter as they both affect demand.

11.3.1.2 People Flow

The demand study started with collecting data for the population flow to build the demand for travel. Researchers from the University of Wisconsin-Madison published a set of human mobility flow data set (Kang et al. 2020), which tracked cell phone signals of commuters from their originating CT block to the destination CT block starting from March 1st, 2020. They summarized the human flow for the local population and visitors visiting the area daily and weekly. Data sets like this could be crucial in route planning of any kind. Here this data set is the base of the GIS model. The visualization of the population flow is the most important factor in the planning of an AAM system.

Since the original data from the University of Wisconsin Madison covered a total time span of two years and was for the entire United States, it was filtered to the CT blocks within the Aerotropolis Atlanta using a MATLAB code. Average and medium population flows were calculated to use in this study case. Origination-Destination (OD) pairs with a distance of fewer than 5 miles were filtered out from the visualization as the distance is too close for AAM service to be effective (Rimjha et al. 2021).

11.3.1.3 Income Consideration

The income of a specific CT block determines the ability of its residents to afford the AAM service. Due to the high startup cost of the AAM system, the cost of building new infrastructures (vertiports), and the acquisition cost of eVTOL aircraft, the price for using an AAM network would be relatively high compared to traditional ground transportation networks (Rimjha et al. 2021). Therefore, higher-income areas are more likely to spend more money on the AAM service instead of using the cheaper traditional ground transportation networks in the early stage of its implementation. Vertiport locations will be chosen in areas with higher medium household income when planning the AAM network.

To visualize the ability to afford the AAM service, medium household income data for each CT block from the Atlanta Regional Committee was overlaid with the demand as shown in Fig. 11.3. Some tracks have endpoints outside of the Aerotropolis Atlanta boundary due to a part of its associated CT blocks being in the boundary.

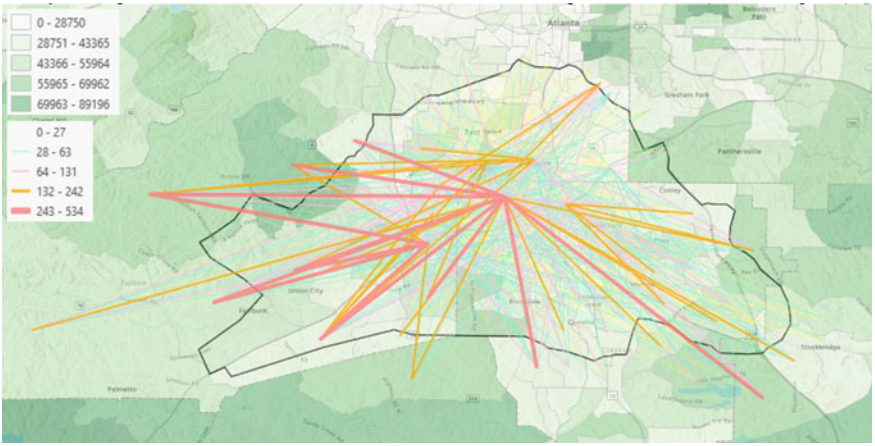


Fig. 11.3 GIS model with population flow and income

11.3.1.4 Existing Public Transportation

Existing public transportation is another factor that would affect the placement of vertiports. The demand for a new and more expensive transportation system may depend on whether there is a successful public transportation system already in place. Competition between the newly established AAM service and the existing public transportation systems would affect the feasibility of the AAM system. To reduce the possibility of competition, existing public transportation networks are modeled using GIS.

Existing bus routes and light rail routes were obtained from Atlanta Regional Committee and modeled in the GIS model. To yield as much demand towards the newly established AAM system from other transportation systems, vertiports in the AAM network are to be located away from the existing stops. The density of the stops is higher in the north of the Aerotropolis Atlanta. Therefore, the factor is considered for more vertiports to be located toward the south of the study area.

11.3.1.5 Airspace Regulations

Airspace regulations are a constraint to the planning process. Complex airspace will drive up the complexity of the problem and could drive up the cost of the process and even make the AAM network infeasible. Within the capability of an eVTOL aircraft, Class B, Class C, and Class D are controlled airspace requiring the aircraft to communicate with the ATC (Designation of Class A, B, C, D, and E 2022). Class E and Class G are uncontrolled airspaces, relatively free to operate in. A Class B airspace is most commonly found surrounding the nation's busiest airports, such as New York's JFK International Airport, San Francisco International Airport, and

Atlanta's Hartsfield Jackson International Air- port. It is the most controlled airspace to provide a safe operating environment for commercial airliners. All traditional aircraft operating in the airspace must receive clearance from the ATC before entering the airspace.

Traditionally, an AAM operation of eVTOL aircraft was considered unfeasible inside a Class B airspace. Due to the strict requirement of obtaining clearance and the massive amount of commercial flights the ATC already has to deal with, adding an AAM network would put too much pressure on the ATC controllers. The increase in the density of aircraft in the airspace would also give the ATC controllers a hard time keeping a safe distance between aircraft.

Based on the FAA UAM ConOps v1.0, eVTOL aircraft and the AAM operation are cleared into controlled airspace inside designated eVTOL corridors without the need to establish communication with or obtain clearance from ATC (Federal Aviation Administration 2020). The document gave greater allowance on the operation as FAA essentially gave the operator of such AAM operation the authority to manage the eVTOL traffic within those established corridors, freeing the planners from the constraint studied in the previously mentioned study.

Even though the UAM ConOps gave the authorization of operating AAM in controlled airspace, some airspace is still to be avoided for all flying activities. Temporary Flight Restriction (TFR) zones are set up temporarily for natural disasters such as wildfires or hurricanes, around largely populated events or locations such as concerts, sporting events or Disney land, or emergency situations. These TFR zones are to ensure public safety during these special events (Special Use Airspace 2022; Federal Aviation Administration 2022b). The only type of TFR zones with concerns are for populated events, as they are most likely to occur in a stadium in a populated city center, where AAM networks are more likely to be set up. Such types of TFR zones are predictable and are designated on the sectional charts for aviation navigation. Therefore the interference to AAM networks can be minimized by planning vertiport locations and routes away from the TFR zones. In this use case, there is no TFR zone that will impact the route planning process. However, in the later use case, the Mercedes-Benz Stadium TFR zone in downtown Atlanta is built in the GIS model for planning the AAM network as a constraint.

11.3.1.6 Noise

Noise regulations are stickily enforced by the FAA, making it one of the more important constraints in the method. The FAA regulates noise emissions from all aviation-related activities. Aviation Noise Abatement Policy (ANAP) issued in 1976 specified noise level above 65 A-weighted decibels (dBA) at residential areas as significant noise exposure and above 75 dBA as severe noise exposure. If the day-night average sound level (DNL) of any residential, school, or medical area exceeds 65 dBA, then the resident or business is eligible for sound insulation paid by the passenger-carrying facility (Federal Aviation Administration 2022c).

In order to avoid exceeding the DNL 65 dBA limit, current airports either plan the arrival and departure routes away from residential areas or reduce night operations as the penalty for exceeding 65 dBA during nighttime is higher when calculating DNL. To avoid exceeding DNL 65 dBA for the AAM network, vertiports are to be located away from residential areas or the arrival and departure routes because the eVTOL aircraft have to be carefully planned to avoid flying over residential areas at low altitudes.

To visualize the location of vertiports with respect to residential areas, land parcel data is required. By overlaying the land use data with the travel demand, vertiport locations can be carefully chosen to satisfy the demand while avoiding high DNL over nearby residential areas. If planners are able to access noise using tools such as the FAA Aviation Environmental Design Tool (AEDT), noise contour plots can be generated (Federal Aviation Administration 2022d). By overlaying the contour plot on the GIS model, DNL noise level near the vertiports can be accurately predicted.

To avoid having a high DNL level, vertiports should have a buffer distance from residential areas directly below the arrival and departure routes. This buffer distance ensures that the eVTOL aircraft, when passing directly above the residential areas, have enough altitude so the noise level experienced by them would be lower than the required environmental standard.

To identify the residential areas near potential vertiport locations, land parcel data was obtained from GIS offices for Fulton County, DeKalb County, and Clayton County. The attribute table for the land parcel shapefiles includes the land use code which was used to identify residential areas. To determine the buffer distance required, measured noise level data from eVTOL manufacturers are needed. Joby Aviation had published the noise level of its S4 eVTOL, stating that at 100 meters from the flight path, the noise level is below 65 dBA (Joby Aviation 2022). The planned AAM network has vertiports with a residential area adjacent to its arrival and departure flight paths. A buffer distance of 700 meters was kept between these vertiports locations and the residential areas, enough for the eVTOL aircraft to climb to a 100-meter altitude.

11.3.2 Ensuring Synergies with Existing Airport

Any non-commercial airliners operating close to existing commercial airports have the potential to affect the operation and cause delays. Since the use case is adjacent to the busiest airport in the world, the Atlanta international airport, this factor is of high importance to the route planning of the AAM network.

The German Aerospace Center studied the impact of an AAM system at a close distance to an existing commercial airport (Ahrenhold et al. 2021). The study simulated an AAM service as an airport shuttle with eVTOL aircraft landing in the airport using both existing runways and designated vertiports located inside the terminals. The results showed that the approaching and departing eVTOL aircraft will only affect the scheduled commercial operations if their route is at a small angle to the

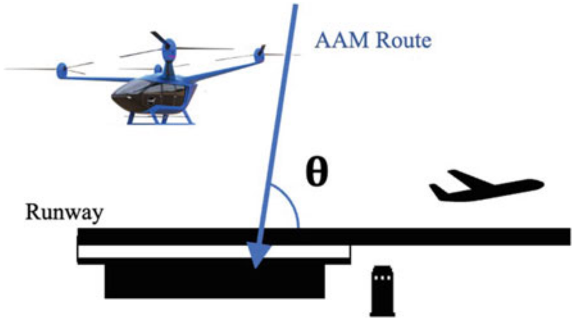


Fig. 11.4 Angle θ between AAM route and airport runway

existing runways. If the eVTOL routes are at a larger angle, or perpendicular to the runway for best results, the AAM network will not affect the commercial air traffic. Figure 11.4 shows an angle θ between the approach route and the runway. This angle θ should be as close to perpendicular or 90 degrees as possible to ensure the synergies.

Therefore, to avoid interference with Atlanta airport traffic, the arrival and departure route to the vertiport located in the airport terminal was planned to be vertical to the runways. The planned AAM network has a route that approaches the Atlanta airport runways perpendicularly.

11.3.2.1 Terrain and Obstacles

Terrain and aviation obstacles (such as hills, antenna towers, and powerlines) are another factor for consideration when planning AAM network routes. Due to the low altitude, the eVTOL aircraft are flying, they are more vulnerable to high obstacles than traditional fixed-wing aircraft or helicopters. To avoid obstacles as much as possible on the route, obstacle data from the aviation sectional chart can be built into the GIS model to aid the route planning process.

Therefore, the last layer in the GIS model shows the obstacles to be avoided for the routes. Aviation obstacles data were provided by Beavers, M. from ESRI. These data were overlaid on top of the commuter demand and income for better visualization to plan the routes.

11.3.2.2 Existing Air Traffic

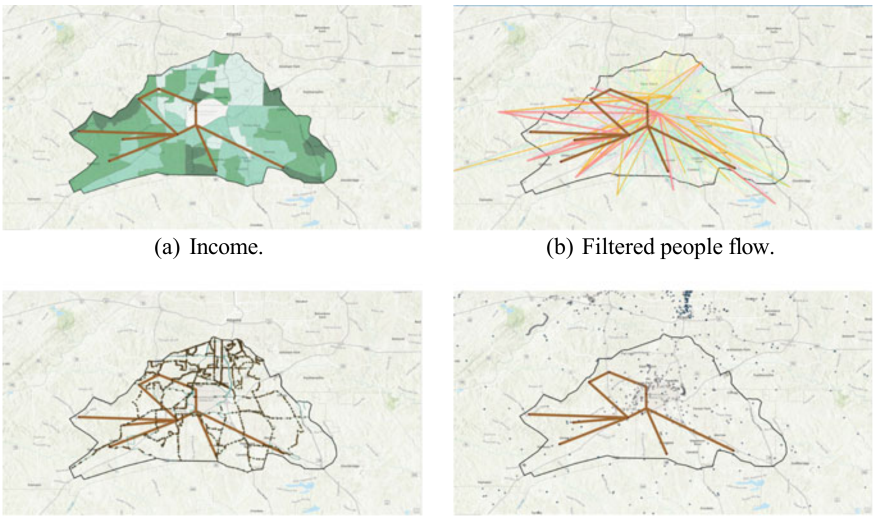
Air traffics such as banner towers and sightseeing helicopters were often seen in downtown Atlanta airspace at extremely low altitude. They are potential collision hazards to eVTOLs during the take-off and approach stages, also at low altitudes. A potential fix to the issue is to establish eVTOL zones around the vertiports with a similar function to an active TFR zone. An eVTOL zone prohibits any aircraft other than an eVTOL to enter, therefore reducing the risk of a mid-air collision between an eVTOL aircraft in the take-off/approach phase and a low-flying traditional aircraft such as a banner tower.

11.3.2.3 Integration of All Factors

With all factors considered, a 9-vertiport AAM network was proposed over the Aerotropolis Atlanta. The demand, existing public transportation networks, and noise factors play important roles in locating the vertiports, while the airport interference and obstacles factors contribute more to the route. Figure 11.5 shows individual layers with the planned AAM network. Figure 11.5a shows the vertiports being placed in higher income areas visualized in darker green. Figure 11.5b shows the vertiports and the route planned following the heaviest people flow. Figure 11.5c shows that the AAM network reaches areas without current public transportation network coverage. Figure 11.5d shows the planned AAM network avoiding higher obstacles in darker blue.

A height profile diagram for the flight path was made to show how the planned height of the AAM network and the clearance it provides to avoid the obstacles along the route (shown in Fig. 11.6).

Figure 11.7 shows the proposed network that is produced based on consideration of all factors and layers.



(a) Income. (b) Filtered people flow.

(c) Existing transportation network. (d) Aviation obstacles.

Fig. 11.5 Layers in the GIS model with the planned AAM network

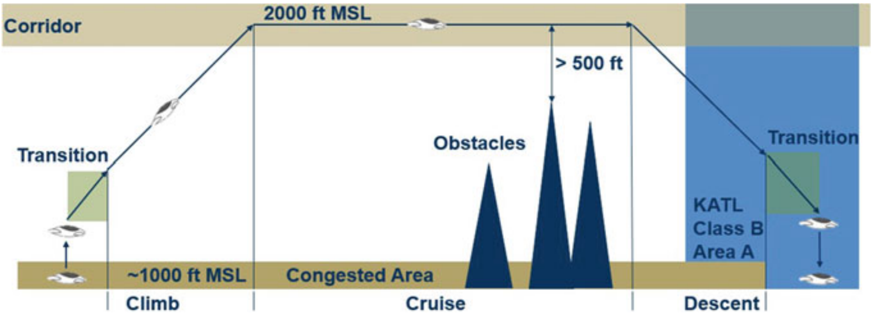


Fig. 11.6 Height profile of the AAM network

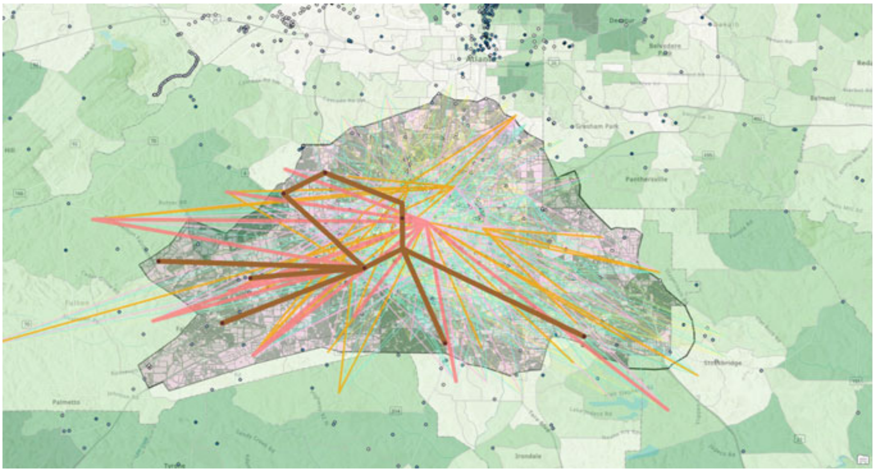


Fig. 11.7 Proposed AAM network in the GIS model

11.3.3 Airport Shuttle Network

The second use case shows route planning for an AAM network connecting the airport to areas of high travel demand. With the concept of remote terminals in mind, the vertiports in this network could reduce the pressure on airport security screening as passengers taking the AAM shuttle could be screened in the vertiports.

11.3.4 Demand

Population flow data from University of Wisconsin Madison was used in this use case too. The data set was filtered to OD pairs with destinations at the Atlanta international airport only. Then the top 20 OD pairs were visualized in 3D GIS using blue cylinders on the originating CT block, as shown in Fig. 11.9. The height of the cylinders was proportional to the number of travelers per day for the OD pair. The top demands originate from locations with high travel demand, including Buckhead, Peachtree City, Newman, and airport hotels adjacent to the airport. Ignoring the demand from airport hotels, there are demands for shuttle services to the other three cities.

The same data from Atlanta Regional Committee was used for the income layer. Shown in Fig. 11.8 is the 3D visualization of the travel demand and medium household income for the area. Both Buckhead, Peachtree City, and Newman have higher medium household incomes compared to their surrounding areas.

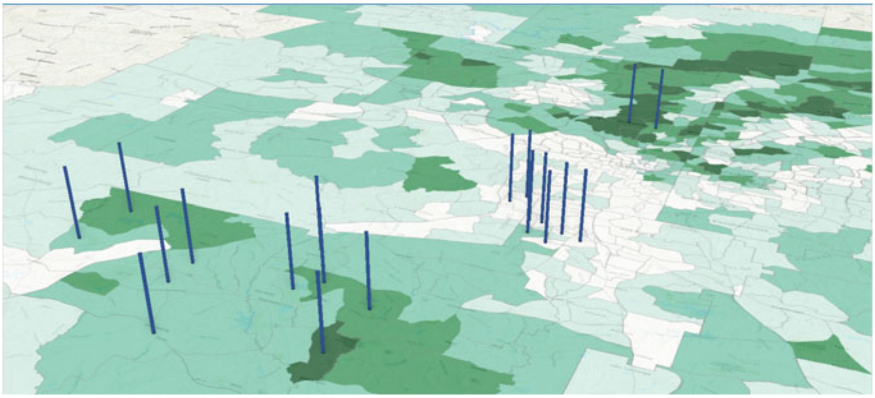


Fig. 11.8 3D demand visualization

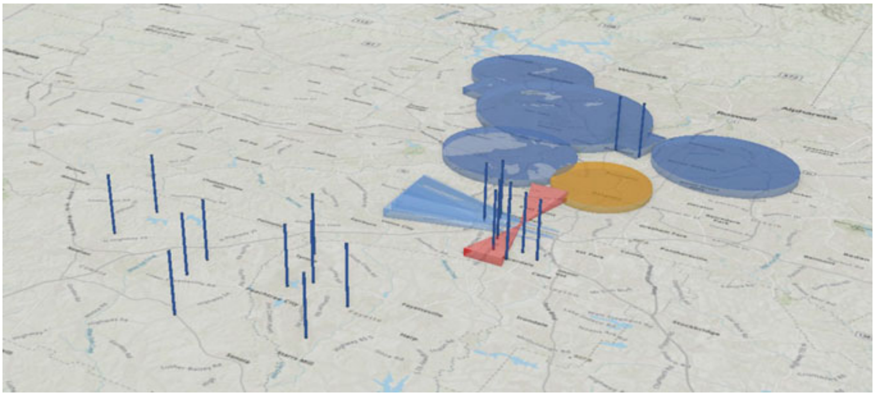


Fig. 11.9 Airspace around atlanta airport in 3D GIS

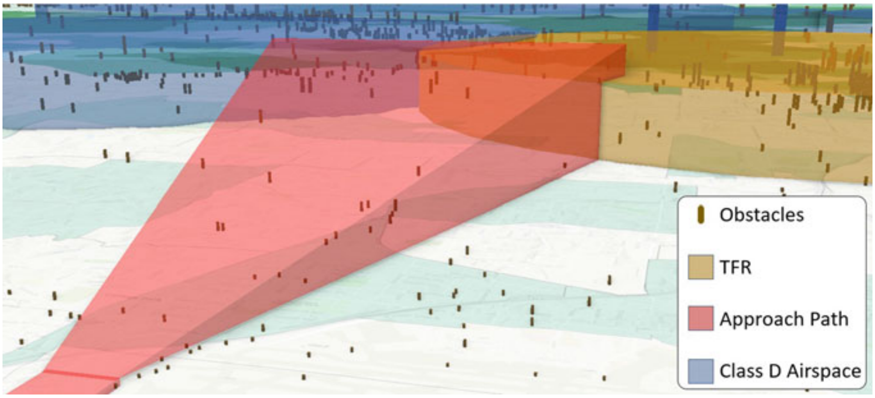


Fig. 11.10 Obstacles around the approach airspace and TFR

11.3.4.1 Airspace and Interference

To visualize the safe airspace for AAM service, a 3D GIS analysis shows a mapping of an airspace associated with airports and the Mercedes-Benz Stadium TFR as shown in Fig. 11.9. Blue cylindrical airspace is class D controlled airspace, and yellow cylindrical airspace is the Mercedes-Benz Stadium TFR. Approach airspace for runway 8 R&L, runway 9R&L, and runway 10 of the Atlanta international airport is also built in the 3D model shown as the blue triangles with their tips pointing to the airport. To avoid interference with international airport traffic, the approach airspace for eVTOLs is built in red to visualize the possible direction and height that the eVTOL aircraft are able to fly.

11.3.4.2 Obstacles

Obstacles are easier to be visualized in the 3D model than in the 2D model. Shown in Fig. 11.10 are the aviation obstacles modeled based on their height. As long as the route stays above the TFR zone and connects to the approach airspace, the obstacles are of no concern for the shuttle network.

11.3.4.3 Integration of Factors

The airport shuttle route can then be planned based on the above geospatial analyses of urban airspace. Originating from Buckhead, Peachtree City, and Newman, the three routes in the shuttle network could be used to connect remote terminals. Figure 11.11 shows the planned routes in the 3D model. Due to the need for high-performance computing power for 3D modeling, the planned network could not be appropriately rendered without the ArcGIS software crashing. Therefore, the current route was

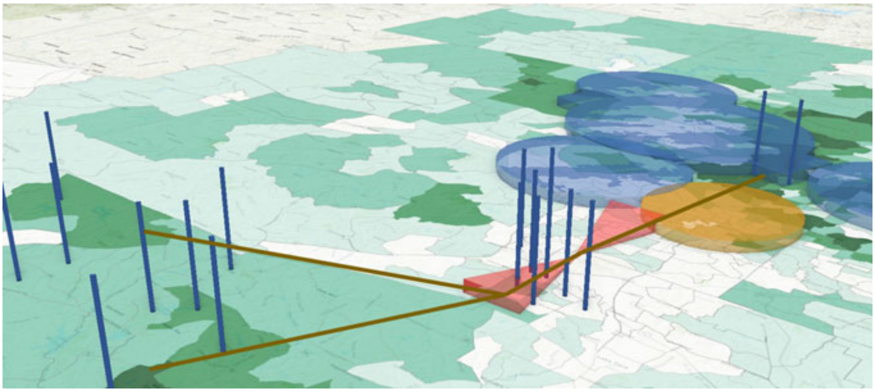


Fig. 11.11 Planned AAM shuttle network

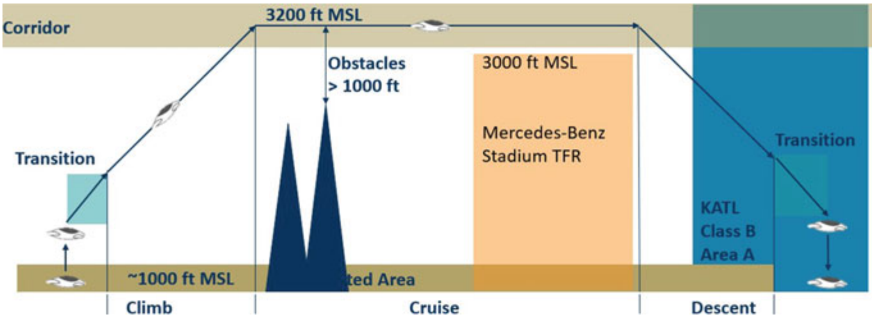


Fig. 11.12 Height profile of the shuttle network

overlaid on top of the ArcGIS model using graphical processing software. Also due to the lack of land use data, vertiport locations can not be determined accurately to avoid exposing residential areas to high noise levels. When the data becomes available, the route planning process can be improved by incorporating more factors.

The height profile diagram for the shuttle network (Fig. 11.12 is made to show the required altitude to clear the Mercedes-Benz Stadium TFR. By flying higher than the TFR, obstacles along the way are also cleared.

11.4 Conclusions and Recommendation

11.4.1 Conclusion

The chapter proposed a method to select the vertiport sites and plan the routes for an AAM system using a GIS environment. By considering factors affecting the network, from FAA regulations to competitions between existing public transportation networks, the method visualizes those factors in the GIS environment to support the decision-making process. Using the Atlanta area for two use cases, different kinds of data needed to account for those factors are introduced and tested. The use case of Inter Aerotropolis Network planned a short-distance inner-city AAM with the purpose of serving as a public transportation network. In contrast, the use case of Airport Shuttle Network is a remote terminal shuttle service for the Atlanta international airport.

11.4.2 Future Work

For future work, it is recommended to develop methods for validating the feasibility of the proposed AAM network and for optimizing the network that aims to be the most effective.

To validate the feasibility of a planned AAM network, Agent-Based Modeling (ABM) could be performed using the shapefiles created. By modeling the demand and all existing traffic networks, including private vehicles and public transportation networks, a baseline case can be simulated. Possible outputs can be an average trip time in the study area and total emissions to complete all trips. Vehicle parameters such as travel speed, acceleration profile, and emission per unit distance traveled are needed in the model to output such results. Then the planned AAM network may be added in the simulation using specific or multiple eVTOL aircraft parameters to compare the results. Since the model has a stochastic behavior, multiple runs are required for each scenario to average the results to ensure accuracy. An example study was done by the Georgia Institute of Technology where the proposed innercity network was simulated in agent-based modeling software to verify its effectiveness in reducing traffic congestion and carbon emission (Wang et al. 2023).

Augmentations can be made to the AAM network, such as increasing the number of vertiports or changing the type of eVTOL aircraft. Future work may include rerunning the ABM model with different inputs using a Design of Experiment (DOE) matrix, then performing a regression analysis to obtain the response surface. With the response surface obtained, optimization can be performed to find the AAM network that best satisfies the requirement of the planners.

References

Airport Noise Compatibility Planning (2022) 14 C.F.R. § 150

Atlanta Aerotropolis (2016) The Aerotropolis Atlanta Blueprint. https://aeroatl.org/wp-content/uploads/2017/06/aerotropolis-atlanta-blueprint-final.pdf. Accessed 15 Oct 2022

Ahrenhold N, Pohling O, Schier-Morgenthal S (2021) Impact of air taxis on air traffic in the vicinity of airports. J Inf. https://doi.org/10.3390/infrastructures6100140

American Society for an Energy Efficient Economy (2020) State and local policy database. https://database.aceee.org/city/atlantaga. Accessed 08 Sep 2022
Bogaisky J (2022) Delta funds electric air taxi maker joby to quicken travel to airports. In: Forbes. Available via Forbes. https://www.forbes.com/sites/jeremybogaisky/2022/10/12/delta-joby-air-taxi-new-york-los-angeles/?sh=69a949052fa0. Accessed 09 Sep 2022
Chitale Y, Introne S, Sotiropoulos E (2021) ConOps Evaluation with scenario integration for urban air mobility. Georgia Institute of Technology
Designation Of Class A, B, C, D, And E (2022) Airspace areas; air traffic service routes; and reporting points, 14 C.F.R. § 71
Federal Aviation Administration (2020) UAM CONOPS v1.0. https://nari.arc.nasa.gov/sites/default/files/attachments/UAMConOps.v1.0.pdf. Accessed 12 Aug 2022
Federal Aviation Administration (2022a) Urban Air Mobility and Advanced Air Mobility. https://www.faa.gov/uas/advancedoperations/urbanairmobility. Accessed 22 Aug 2022
Federal Aviation Administration (2022b) Temporary Flight Restrictions (TFRs). https://www.faa.gov/uas/gettingstarted/temporaryflightrestrictions. Accessed 30 Sep 2022
Federal Aviation Administration (2022c) FAA History of Noise. https://www.faa.gov/regulationspolicies/policyguidance/noise/history. Accessed 28 Sep 2022
Federal Aviation Administration (2022d) Aviation Environmental Design Tool (AEDT) Version 3e User Manual. Washington D.C.
INRIX (2021) Global traffic scoreboard. https://inrix.com/scorecard/. Accessed 08 Sep 2022
Joby Aviation (2022) Joby confirms revolutionary low noise foot-print following NASA testing. https://www.jobyaviation.com/news/joby-revolutionary-low-noise-footprint-nasa-testing/. Accessed 22 Sep 2022
Kang Y, Gao S, Liang Y, et al (2020) Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic. J Sci Dat. https://doi.org/10.1038/s41597-020-00734-5
Kasarda D (2013) Airport cities: the evolution. In: Airport world magazine, pp 24–27. http://aerotropolisbusinessconcepts.aero/wp-content/uploads/2014/08/3-AirportCities.TheEvolution2.pdf. Accessed 17 Jan 2023
Rimjha M, Hotle S, Trani A et al (2021) Commuter demand estimation and feasibility assessment for urban air mobility in Northern California. J Tra Res. https://doi.org/10.1016/j.tra.2021.03.020
Special Use Airspace (2022) 14 C.F.R. § 73
Vascik P, Cho J, Bulusu V et al (2020) Geometric approach towards airspace assessment for emerging operations. J Air Transp 28(3):124–133
Vehicle Technology Office (2015) Idling reduction for personal vehicles. U.S. Department of Energy. https://afdc.energy.gov/files/u/publication/idlingpersonal-vehicles.pdf. Accessed 08 Sep 2022
Wang X, Balchanos M, Mavris D (2023) A feasibility study for the development of air mobility operations within an airport city (Aerotropolis). Paper presented at the AIAA SciTech 2023 forum, National Harbor, 22–27 Jan 2023

Part III
Fine-Scale Urban Analysis

Chapter 12
"Eyes on the Street": Estimating Natural Surveillance Along Amsterdam's City Streets Using Street-Level Imagery

Timo Van Asten, Vasileios Milias, Alessandro Bozzon, and Achilleas Psyllidis

Abstract Neighborhood safety and its perception are important determinants of citizens' health and well-being. Contemporary urban design guidelines often advocate urban forms that encourage natural surveillance or "eyes on the street" to promote community safety. However, assessing a neighborhood's level of natural surveillance is challenging due to its subjective nature and a lack of relevant data. We propose a method for measuring natural surveillance at scale by employing a combination of street-level imagery and computer vision techniques. We detect windows on building facades and calculate sightlines from the street level and surrounding buildings across forty neighborhoods in Amsterdam, the Netherlands. By correlating our measurements with the city's Safety Index, we also validate how our method can be used as an estimator of neighborhood safety. We show how perceived safety varies with window level and building distance from the street, and we find a non-linear relationship between natural surveillance and (perceived) safety.

Keywords Natural surveillance · Perceived safety · Crime · Eyes on the street · Street-level imagery

---

T. Van Asten
Dutch Ministry of the Interior and Kingdom Relations, The Hague, The Netherlands

V. Milias (✉) · A. Bozzon · A. Psyllidis
Faculty of Industrial Design Engineering, Delft University of Technology, Landbergstraat 15, 2628CE Delft, The Netherlands
e-mail: v.milias@tudelft.nl

A. Bozzon
e-mail: a.bozzon@tudelft.nl

A. Psyllidis
e-mail: a.psyllidis@tudelft.nl

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_12

12.1 Introduction

Safe outdoor environments contribute substantially to a neighborhood's level of livability. A growing body of literature has shown that the design and structure of the built environment can influence both actual safety risks and how safety is perceived by different population groups, subsequently impacting citizens' physical and mental health and well-being (Foster and Giles-Corti 2008; Jackson and Stafford 2009; Mason et al. 2013; Stafford et al. 2007). Especially in neighborhoods where the fear of crime is disproportionate to the actual crime rates, there is evidence of significant associations with lower levels of physical activity (e.g., limited play among children and walking in older populations, or women being discouraged from using parts of the neighborhood), leading to increased levels of childhood obesity and social isolation amongst the elderly (Barnett et al. 2017; Groshong et al. 2020; Won et al. 2016).

Design and planning approaches to community safety in urban spaces, including Newman's defensible space theory (Newman 1972) and the Crime Prevention Through Environmental Design (CPTED) strategic framework (Crowe 2000), often advocate for neighborhoods with increased density, mixture of land uses, well-maintained walkways, and permeable street networks with high connectivity, even though there has been some recent criticism about the universality of these features in reducing actual crime risk and the fear of crime (Barton 2010; Cozens 2015; Cozens and Hillier 2012; Rydin et al. 2012). Similar principles are also adopted by the recent United Nations' guidelines on safer cities and human settlements (UN-Habitat 2020).

A common denominator across theories and design approaches aimed at improving actual and perceived safety is the provision of *natural* (or *passive*) *surveillance* in urban public spaces (Crowe 2000). Originating in what Jane Jacobs referred to as "eyes on the street" (Jacobs 1961), enhancing a neighborhood's level of natural surveillance has become a widely adopted design guideline towards safer urban environments (Carmona 2021; UN-Habitat 2020). Natural surveillance is a byproduct of how citizens normally and routinely use public spaces (Crowe 2000). Even though there are several factors that can influence the level of natural surveillance, it is generally assumed that characteristics such as good street lighting, abundance of unobstructed windows overlooking walkways, and more permeable streets contribute to an increased level of natural surveillance (Cozens 2015; Foster et al. 2011). However, evidence of how much of these built-environment characteristics contribute to increased natural surveillance and lead to improved perceptions of safety is still lacking. Several approaches to measuring natural surveillance have been proposed to date, ranging from collecting observations on the ground (Lee et al. 2017; Peeters and Vander Beken, 2017; Reynald and Elffers 2009) to computational models for estimating sightlines (Amiri and Crain 2020; Shach-Pinsly 2019). Yet, measurements across large spatial extents remain a challenge even for computational approaches, primarily due to the subjective nature of surveillance and the lack of relevant fine-grained data.

This paper addresses these knowledge gaps, first, by introducing a method for measuring natural surveillance at scale using street-level imagery and computer vision techniques and, second, by providing evidence of a non-linear relationship between natural surveillance and (perceived) safety. We collected and analyzed street- level imagery along all street segments across 40 neighborhoods in the city of Amsterdam, the Netherlands. Unlike related approaches that use generic proxies of visibility such as the distance between buildings (De Nadai et al. 2020; Shach-Pinsly 2019), street-level imagery gives us the opportunity to capture built-environment features that can affect natural surveillance, such as the location of windows on a building facade and any visibility blockages by fences or vegetation.

We extracted these features with geolocalization and computer vision (i.e., facade labeling) techniques. We calculated sightlines from the windows to each street and vice versa, as well as the windows of surrounding buildings, using the extracted features. Drawing on the work of (Peeters and Vander Beken 2017) and (Amiri and Crain 2020), we calculated two types of surveillance for each street segment: (1) *street surveillance*, which captures the surveillance of windows from the street level and vice versa, and (2) *occupant surveillance*, which captures the surveillance of windows from surrounding buildings. We then correlated the resulting surveillance values per street segment with publicly available data on (perceived) safety, crime, and nuisance in Amsterdam. Our research goes beyond defining the magnitude of associations between natural surveillance and (perceived) safety by identifying street-segment surveillance threshold values above which the feeling of safety remains unchanged. Such evidence can have significant implications for the design of safer neighborhoods and communities.

The remainder of this paper is structured as follows. We first present our research methods and describe the data sources, the study area, and how we calculated the sightlines and measure street and occupant surveillance. We then report the results of our analysis of Amsterdam's neighborhoods, and correlate them with the Amsterdam Safety Index. Next, we discuss the outcomes of our analysis, identify threshold values and implications for the design of safer communities. Finally, we summarize the conclusions and suggest future lines of research.

12.2 Method

We estimated natural surveillance at the street-segment level considering both *street* and *occupant* surveillance[1]. Both of these surveillance types depend on the degree to which people are able to observe the street from a specified distance; what is usually referred to as *sightline*. We grouped sightlines according to three parameters. The first parameter was based on the assumption that surveillance from ground floor windows is associated with lower levels of street crime than surveillance from up-

[1] The source code of our method is available on GitHub: https://github.com/timovanasten/natural-surveillance.

per floor windows, which is supported by previous research (Lee et al. 2017). We grouped sightlines based on their altitude $a_{\max}$ to study the impact surveillance from different window levels has on street safety. Assuming that each building story is approximately 3 m high, the value of $a_{\max}$ for first-floor windows is 3 m, for first and second-floor windows it is 6 m, and for first, second, and third-floor windows it is 9 m. Second, existing literature indicates that the most reliable distance to observe and interpret an event is 15 m (Amiri and Crain 2020; Jong et al. 2005; Lindsay et al. 2008). Furthermore, events witnessed from a 43-m distance produce weak but reliable eyewitness accounts. Therefore, we divided sightlines into two types: those with $d_{\max} \leq 15$ m and those with $d_{\max} \leq 43$ m. Finally, we defined an angle θ_{fov} as the field of view visible through a window. Outside of this field of view, sightlines were excluded.

Detection and geolocalization of windows. The first step in the calculation of sightlines is the detection of windows on building facades along streets. We collected street-level imagery along the streets of interest and detected windows using the facade labeling algorithm developed by Li et al. (2020). The algorithm detects a set of four key-points (i.e., top-left, bottom-left, bottom-right, and top-right) using 2D heatmaps. Then, it links them together using a neural network trained on labeled images with varying facade structures, viewing angles, lighting, and occlusion conditions. Following this, we calculated the geolocation of each detected window by using and adapting the geolocalization algorithm that was originally developed in Qiu et al. (2019) and later modified in Sharifi Noorian et al. (2020). For each of the detected windows, this process calculates the latitude, longitude, and altitude of the four key-points in the street-level images (Fig. 12.1). We then computed the center-point and stored it as the window's geolocation.

Street and occupant surveillance. To calculate *street surveillance*, we counted all sightlines that have as a starting point the windows with an altitude lower than a_{max} and as an endpoint the location of the street-car camera, with a length lower than $d_{\max}$. The number of these sightlines reflects the number of windows that have unobstructed views to points sampled across the street network (i.e., every 10 m). Regarding *occupant surveillance*, we calculated for each window the number of

Fig. 12.1 Indicative outputs of the facade labeling algorithm developed by Li et al. (2020) on street-level images collected in Amsterdam

neighboring windows that have an unobstructed view (i.e., sightline) to it. Specifically, for each detected window $o_{\text{viewpoint}}$ with an altitude lower than $a_{\max}$, we selected all neighboring windows that had an altitude lower than $a_{\max}$ and were at a maximum distance $d_{\max}$ from each $o_{\text{viewpoint}}$. Next, we calculated all sightlines from the neighboring windows to the $o_{\text{viewpoint}}$ and removed the ones that were obstructed by the presence of intermediate buildings. In particular, for each sightline s, we calculated the angle θ_s between s and the building segment that contains $o_{\text{viewpoint}}$. If $\theta_s > \frac{1}{2}\theta_{fov}$, we considered s outside the field of view of the neighboring window and removed it from the set of sightlines to be considered. Due to the restriction of the sightline angle, a sightline originating from each window ($o_{\text{viewpoint}}$) to each neighboring window (o_{neighbor}) does not imply that the reverse also exists. By repeating these steps for each detected window, we calculated how many neighboring windows have unobstructed views of each window at hand.

To calculate the overall *natural surveillance* scores, we linked all points with a street and occupant surveillance score to a given street segment q. We defined a street segment as a section of the street between two junctions, or between a junction and the end of the street, if the street has a dead end. More specifically, each window was linked to the image where it was detected, and this is, in turn, linked to the corresponding street segment. We calculated the following two scores, normalized by the street segment's length:

$$S_q = \frac{1}{qL} \sum_{i \in P_q} s_i \tag{12.1}$$

$$O_q = \frac{1}{qL} \sum_{i \in P_q} o_i \tag{12.2}$$

where S_q and O_q respectively denote the street and occupant surveillance scores, P_q denotes the set of points linked to a street segment q, q_L is the length of street segment q in meters, and s_i and o_i denote the number of sightlines observing point i.

We further aggregated our scores at the neighborhood level by calculating the sum of the sightlines of all points within each neighborhood, and dividing it by the length of each street segment using the following formulas:

$$S_n = \frac{\sum_{i \in P_n} s_i}{\sum_{i \in Q_n} qL_i} \tag{12.3}$$

$$P_n = \frac{\sum_{i \in P_n} o_i}{\sum_{i \in Q_n} qL_i} \tag{12.4}$$

where S_n and O_n respectively denote the sum of street and occupant surveillance scores within a neighborhood n, Q_n is the set of sampled street segments within n, and P_n is the set of sampled points linked to the street segments within n. As previously stated, we further grouped our scores according to whether the distance between the windows and the corresponding street segment was reliable for witnessing an event ($d_{max} = 15\ m$) or dependable ($d_{max} = 43\ m$).

Correlation analysis. To investigate the relationship between natural surveillance and safety (both actual and perceived), we correlated the street surveillance and occupant surveillance scores with the Amsterdam Safety Index (Gemeente Amsterdam 2015). We first tested the normality of the estimated natural surveillance scores using the Kolmogorov–Smirnov test, which indicated a non-normal distribution. Therefore, for each of the considered neighborhoods, we used Spearman's rank correlation coefficient (*rho*) to calculate the correlation between the natural surveillance scores and the Index's aggregate values and sub-components (i.e., crime, nuisance, and perceived safety).

12.3 Data

We used the city of Amsterdam in the Netherlands as a case study to illustrate and validate how our method could be used to assess a neighborhood's level of natural surveillance. Amsterdam is the capital and most populated city in the Netherlands, characterized by a variety of neighborhoods with equally varying levels of reported (perceived) safety. The city combines a medieval center bustling with tourists all year round with new developments and strictly residential areas in the outskirts. A well-substantiated dataset on (perceived) safety for the entire city of Amsterdam is publicly available, making it an exemplary case to compare our measurements with real-world data on actual and perceived safety. The Amsterdam Safety Index (Gemeente Amsterdam 2015) covers 104 neighborhoods in the city and is composed of three sub-components. These are, namely, the levels of *crime*, *nuisance*, and *perceived safety* in each neighborhood. The lower the index value, the safer the neighborhood is considered to be. We used OpenStreetMap (OSM), an open-source mapping platform, to collect data about the street network and the building footprints. We made use of the *OSMnx* (Boeing 2017) Python library to extract street network data, and the OSM *Overpass* Application Programming Interface (API) to collect the building footprints. Moreover, we used the Google Street View Static API to detect the location of windows on the building facades along the street segments. Due to budget limitations, we focused on 40 out of the 104 Amsterdam neighborhoods covered by the Safety Index. Neighborhoods were selected such that they are spatially contiguous (i.e., they share a common administrative boundary) and are characterized by a variety of safety index scores. In the final subset of 40 neighborhoods (Fig. 12.2), we collected 6,667 street segments from OSM and 109,988 street-level images from Google Street View along the street segments, with the maximum allowed resolution of 640×640 pixels and with orthogonal field of view and zero pitch to capture the building facades. The window detection and geo-localization algorithms provided a total of 872,360 windows, extracted with the use of the ResNet18 model for facade labeling.

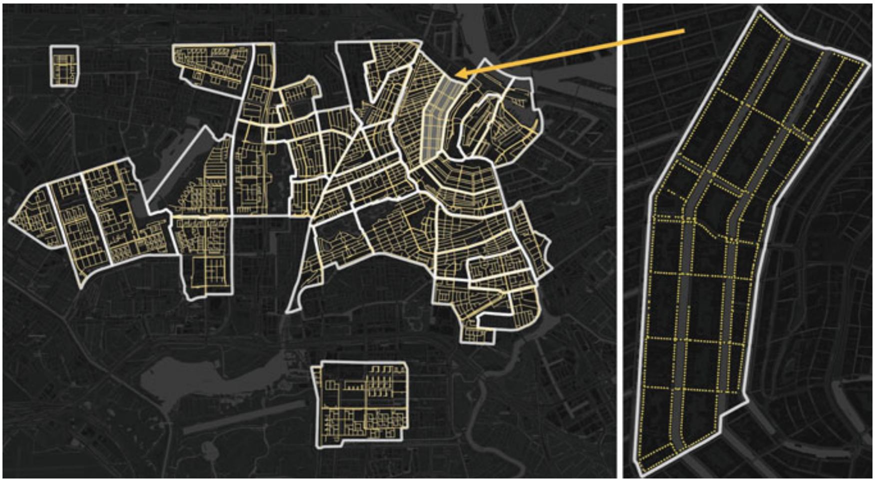


Fig. 12.2 Left: The 40 neighborhoods and their streets considered in our analysis. Right: An example of the geolocations of street-level images collected along street segments in Amsterdam's Grachtengordel-West neighborhood

12.4 Results

This section provides an overview of the application of our method for estimating natural surveillance in Amsterdam, the Netherlands. We also present the correlation results between our measurements of street and occupant surveillance and the 2019 Amsterdam Safety Index and its sub-components, namely crime, nuisance, and perceived safety in each of the considered neighborhoods. Furthermore, we compared the influence of considering windows from different floor levels and distances from the street on indicating a neighborhood's actual and perceived safety levels.

Figure 12.3 illustrates the calculated street and occupant surveillance scores, considering windows within a distance of 43 m from the streets and up to the first floor of the buildings, together with the overall Safety Index values of the 40 considered neighborhoods. We converted each of these scores into three categorical variables, namely, low, medium, and high, according to the tertile they belong to. This allows for easier visual comparison, given that each of the presented metrics is originally expressed in different units.

Most neighborhoods showcased consistency across the three scores, routinely resulting in high or low safety areas. Indicative examples of this include the Burgwallen- Oude Zijde (BOZ) and Burgwallen-Nieuwe Zijde (BNZ) neighborhoods, both in the historical center of Amsterdam, consistently scoring low across all metrics. Similarly, neighborhoods such as Museumkwartier (MU) and Staatsliedenbuurt (ST) consistently scored among the safest. Examples of the opposite include Buitenveldert-West (BW), which scored low in terms of street surveillance, medium in occupant surveillance, and high in the Safety Index values. Figure 12.3 further

Fig. 12.3 **a** Street surveillance scores; **b** Occupant surveillance scores; **c** Amsterdam Safety Index values for the 40 considered Amsterdam neighborhoods, classified into tertiles

zooms in on the street structure of select neighborhoods to elucidate the individual contributions of street segments to the overall street and occupant surveillance scores.

Table 12.1 shows the results of the correlation between our street and occupant surveillance scores and the Amsterdam Safety Index and its sub-components, using the Spearman's rank correlation coefficient (*rho*). The results yielded a moderate negative and statistically significant correlation ($r = -0.49$, $p < 0.001$) between street surveillance and Safety Index values in the case of sightlines with $d_{\max} = 43$ m and up to the first floor of buildings (i.e., 1F). The correlation became weaker when we considered sightlines from buildings within a 15 m distance, or from higher floors. Looking at the Index's sub-components, the correlation between street surveillance and *crime* or *nuisance* also became weaker when we considered sightlines originating from floors higher than the first. Also, street surveillance scores generally presented strong negative correlations with *perceived safety* values, with sightlines of 15-m length yielding the strongest results. The correlations of occupant surveillance scores with the Safety Index and its sub-components were generally weaker in comparison with their street surveillance counterparts. The occupant surveillance scores had the highest correlation with the average Safety Index values ($r = 0.34$).

The use of a locally weighted scatter plot smoothing (LOWESS) regression to examine the linearity of the relationship between the different scores, as shown in Fig. 12.4, provided additional insight into the correlations. Specifically, the comparison of the street surveillance scores with the overall Safety Index (Panel I) and the perceived safety sub-component (Panel II) yielded an interesting non-linear pattern.

Table 12.1 Spearman correlations of the street and occupant surveillance scores with the 2019 Amsterdam Safety Index

Correlations of street and occupant surveillance with the Amsterdam Safety Index							
Index and sub-components	Maximum sightline length						
	15 m (reliable)			43 m (dependable)			None
	Included floors						
	1F	≤2F	≤3F	1F	≤ 2F	≤ 3F	All
	Street surveillance						
Safety Index	−0.45 **	−0.38*	−0.39*	**−0.49****	−0.42**	−0.40*	−0.45**
Crime	−0.34*	−0.30	−0.31	**−0.40***	−0.37*	−0.34*	−0.33*
Nuisance	−0.23	−0.18	−0.15	**−0.29**	−0.21	−0.18	−0.27
Perceived Safety	**−0.48****	−0.43**	−0.44**	−0.45**	−0.36*	−0.38*	−0.43**
	Occupant surveillance						
Safety Index	−0.34*	−0.38*	−0.39*	−0.40*	**−0.41***	−0.38*	
Crime	−0.23	−0.33*	−0.34*	−0.31*	**−0.35***	−0.31	
Nuisance	−0.25	−0.22	−0.20	**−0.32***	−0.24	−0.19	
Perceived Safety	**−0.36***	−0.35*	**−0.36***	−0.32**	−0.33*	−0.35*	

$* = p <= 0.05$, $** = p <= 0.01$

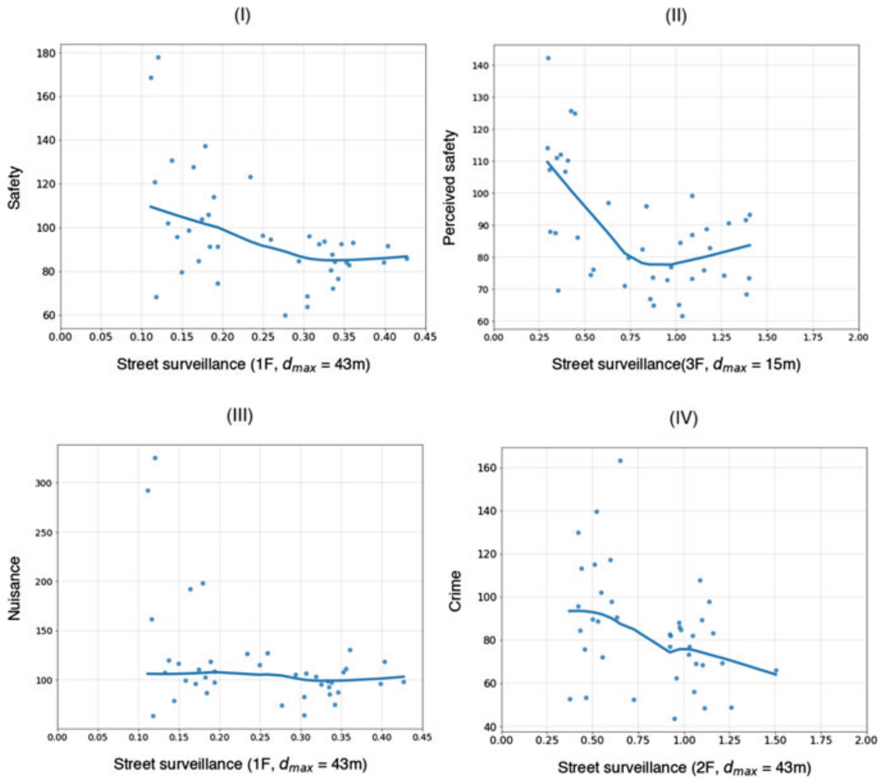


Fig. 12.4 Scatter plots and corresponding trendlines depicting the relationship between estimated street surveillance scores and the Amsterdam Safety Index (I) and its sub-components: perceived safety (II); nuisance (III); crime (IV)

We observed that as the street surveillance score increased, the trendline became relatively horizontal from a certain point onward. This suggests that even though an increased level of street surveillance is generally associated with higher safety—either actual or perceived—this association becomes weaker after a certain value (approximately 0.3 for safety and 0.8 for perceived safety). However, this does not seem to be the case when it comes to the association of street surveillance with the levels of nuisance (Panel III) and crime (Panel IV). The corresponding scatter plots and trendlines did not indicate any particular pattern.

12.5 Discussion

The application of our method in several neighborhoods in Amsterdam demonstrated that the combination of street-level imagery with computer vision techniques offers a promising approach to measuring natural surveillance across large spatial extents. As such, it can provide a pathway to integrate the widely advocated, yet difficult to capture, notion of “eyes on the street” into the planning and design for safer neighborhoods. We also showed that different aspects of surveillance (i.e., street or occupant-based) have varying contributions to a neighborhood’s overall level of safety and elicit interesting non-linear relationships.

Our results suggest an overall significant negative correlation between our average natural surveillance scores and the Safety Index. Given that lower Index values indicated safer neighborhoods, an increased level of natural surveillance correspondingly indicated a safer neighborhood. This aligns with related expectations from the CPTED literature (Cozens and Love 2015; Crowe 2000). We specifically detected stronger correlations between street surveillance scores and the Safety Index, whereas the correlations with occupant surveillance scores appeared weaker. This suggests that the degree of street visibility from surrounding windows is a stronger predictor of a street’s level of natural surveillance compared to window visibility from surrounding buildings.

Our analysis also uncovers aspects of natural surveillance that have largely been overlooked in the existing evidence base. Our findings, in particular, support our hypothesis that window levels and the distance of buildings from the street influence how natural surveillance correlates with overall safety. However, variations do exist between the two different facets of natural surveillance. Specifically, street visibility (i.e., street surveillance) from first-floor windows correlates most with the average values of the Safety Index and this association becomes stronger as the distance from the street increases. This appears to be less the case when it comes to the visibility of windows from surrounding buildings (i.e., occupant surveillance). We also observed a strong correlation between street surveillance and perceived safety. The consideration of different floor levels barely influenced this correlation. However, the distance of the windows from the street did influence this association, with windows closer to the street (i.e., sightlines with a length of 15 m) leading to stronger correlations.

The overall strong association between street surveillance and perceived safety would generally suggest that the higher the level of street surveillance, the more it would correlate with an increased perception of safety. However, our analysis of linearity indicates a natural surveillance threshold value above which the level of perceived safety remains relatively stable. Figure 12.4 (Panel I) shows that street segments with street surveillance score up to 0.3—this corresponds to an average of three windows overlooking a 10-m-long street segment—accordingly led to a gradually increasing level of neighborhood safety. However, street surveillance scores of 0.3 and above did not lead to any increases in the overall level of safety. Figure 12.5 provides indicative examples of streets with street surveillance scores of 0, 0.15, 0.3, and 0.7 to showcase what streets of low or high street surveillance are like. This

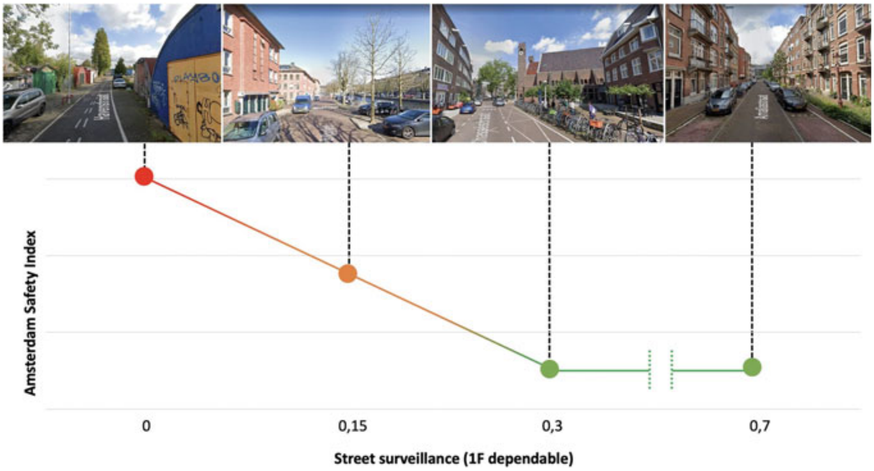


Fig. 12.5 A selection of Amsterdam streets, along with their street surveillance scores (1F reliable). Streets included in the figure, from left to right: Havenstraat, De Rijpgracht, Chasse' straat and Amaliastraat. As the street surveillance score rises to 0.3, neighborhoods feel safer. An increase to 0.7 does not appear to be associated with increased (perceived) safety

result provides an interesting insight into *how much* and *what kinds* of street features lead to increased (feelings of) safety and requires further research in different urban environments.

There are several limitations in this study that could be addressed in future research. First, our method depends on how well windows are depicted in street-level imagery. Some windows, however, are excluded from the street-level images either due to a lack of imagery along certain street segments or because the view from the street to the window is obstructed (e.g., by a tree). Second, the facade labeling algorithm used for window detection is occasionally inaccurate. Indicative inaccuracies include windows that are in shadow or with occluded openings, as well as storefront windows that may go undetected by the algorithm because it was trained on images of residential buildings (Li et al. 2020). Nonetheless, the facade labeling algorithm we used was tested on several datasets of building facade images and achieved a pixel accuracy of 90% on average (Korc and Förstner 2009; Riemenschneider et al. 2012; Teboul et al. 2011). Third, the geolocalization algorithm can introduce errors (namely, mean error of 1.07 m and standard deviation of 1.09 m), the most significant of which are caused by inaccurate GPS metadata in street-level images. Fourth, natural surveillance is a broad concept that encompasses more than window-based surveillance. In fact, active street observation is largely dependent on residents' time, willingness, and capacity to watch and defend their streets and communities (Cozens 2015). Therefore, our research could be expanded to take into account other factors that influence natural surveillance, such as citizens' daily activity patterns, time of day, or street lighting quality. Lastly, we only tested and evaluated our method in a few Amsterdam neighborhoods. For this reason, we intend to broaden the applicability of our method by implementing it in other cities.

12.6 Conclusion

Natural surveillance has drawn a lot of interest and is now a crucial component of design strategies aimed at improving actual and perceived safety in urban spaces. This paper introduced a method for measuring natural surveillance at scale by leveraging a combination of street-level imagery and computer vision techniques. Our work has practical value for built-environment professionals who seek to understand and improve the levels of actual and perceived safety in urban neighborhoods. Specifically, our method draws on the new possibilities offered by street-level imagery in capturing built-environment features, such as the location of windows and any visibility blockages that have been shown to affect natural surveillance. It also employs geolocalization and facade labeling techniques to estimate the surveillance of windows from the street level and surrounding buildings across large spatial extents. We applied our method in neighborhoods of Amsterdam and correlated our measurements with various components comprising the city’s Safety Index to validate its use as an estimator of neighborhood safety. Our analysis showcased that our method can be a promising and scalable alternative to the manual collection of observations around aspects of natural surveillance. Our results align with existing evidence from the CPTED literature and suggest that surveillance of windows from the street contributes more to the overall safety than surrounding buildings. Moreover, the window level and distance of buildings from the street appear to have varying influences on the feeling of safety. Another intriguing finding from our analysis is the identification of a natural surveillance threshold point, with scores above it not resulting in increased (perceived) safety. However, more research in different urban settings is required to provide more evidence of this. In particular, it presents an interesting avenue of future research in the fields of environmental criminology and next-generation CPTED (Saville and Cleveland 2008) that can contribute to an improved understanding of what kinds of built-environment characteristics lead to increased (feelings of) safety.

Acknowledgements Vasileios Milias, Alessandro Bozzon, and Achilleas Psyllidis acknowledge support from the European Union’s Horizon 2020 research and innovation program under grant agreement No. 874724.

Author Contribution Statement Timo Van Asten and Vasileios Milias equally contributed to the study conception, research design, methodology, data analysis, and writing of the draft and revised manuscripts. Alessandro Bozzon contributed to the research design, funding acquisition, resources, supervision, review, and editing of the manuscript. Achilleas Psyllidis contributed to the study conception, research design, methodology, funding acquisition, resources, supervision, and writing of the draft and revised manuscripts.

References

Amiri S, Crain DR (2020) Quantifying jacobs' notion of 'eyes upon the street' in 3-dimensions. J Urban Des 25(4):467–485. https://doi.org/10.1080/13574809.2019.1691440,URL

Barnett DW, Barnett A, Nathan A, Van Cauwenberg J, Cerin E (2017) Built environmental correlates of older adults' total physical activity and walking: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 14(1):1–24

Barton H (2010) Strengthening the roots of planning. Plan Theory Pract 11(1):95–101

Boeing G (2017) Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput Environ Urban Syst 65:126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004, http://arxiv.org/pdf/1611.01890

Carmona M (2021) Public places urban spaces: the dimensions of urban design. Routledge

Cozens P, Hillier D (2012) Revisiting Jane Jacobs's 'eyes on the street' for the twenty- first century: evidence from environmental criminology, Routledge, pp 196–214. https://doi.org/10.4324/9780203095171

Cozens P, Love T (2015) A review and current status of crime prevention through environmental design (cpted). J Plan Lit 30(4):393–412

Cozens P (2015) Crime and community safety: challenging the design consensus. In: The Routledge handbook of planning for health and well-being, Routledge, pp 196–211

Crowe TD (2000) Crime prevention through environmental design: applications of architectural design and space management concepts, 2nd edn. Butterworth- Heinemann, Boston, Mass

Foster S, Giles-Corti B (2008) The built environment, neighborhood crime and constrained physical activity: an exploration of inconsistent findings. Prev Med 47(3):241–251

Foster S, Giles-Corti B, Knuiman M (2011) Creating safe walkable streetscapes: does house design and upkeep discourage incivilities in suburban neigh-bourhoods? J Environ Psychol 31(1):79–88. https://doi.org/10.1016/j.jenvp.2010.03.005

Gemeente Amsterdam (2015) Verantwoordingsdocument veiligheidsindex 2014. Report

Groshong L, Wilhelm Stanis SA, Kaczynski AT, Hipp JA (2020) Attitudes about perceived park safety among residents in low-income and high minority kansas city, missouri, neighborhoods. Environ Behav 52(6):639–665

Jackson J, Stafford M (2009) Public health and fear of crime: a prospective cohort study. Br J Criminol 49(6):832–847

Jacobs J (1961) The death and life of great American cities

Jong MD, Wagenaar WA, Wolters G, Verstijnen IM (2005) Familiar face recognition as a function of distance and illumination: a practical tool for use in the courtroom. Psychol Crime Law 11(1):87–97. https://doi.org/10.1080/10683160410001715123

Korc F, Förstner W (2009) eTRIMS image database for interpreting images of man-made scenes. Dept of Photogrammetry, University of Bonn, Tech Rep TR-IGG-P-2009-01

Lee I, Jung S, Lee J, Macdonald E (2017) Street crime prediction model based on the physical characteristics of a streetscape: analysis of streets in low-rise housing areas in south korea. Environ Plann B Urban Analytics City Sci 46(5):862–879. https://doi.org/10.1177/2399808317735105

Li CK, Zhang HX, Liu JX, Zhang YQ, Zou SC, Fang YT (2020) Window detection in facades using heatmap fusion. J Comput Sci Technol 35(4):900–912. https://doi.org/10.1007/s11390-020-0253-4, http://jcst.ict.ac.cn/CN/abstract/article2660.shtml

Lindsay RC, Semmler C, Weber N, Brewer N, Lindsay MR (2008) How variations in distance affect eyewitness reports and identification accuracy. Law Hum Behav 32(6):526–535. https://doi.org/10.1007/s10979-008-9128-x

Mason P, Kearns A, Livingston M (2013) "safe going": the influence of crime rates and perceived crime and safety on walking in deprived neighbourhoods. Soc Sci Med 91:15–24

De Nadai M, Xu Y, Letouze´ E, Gonza´lez MC, Lepri B (2020) Socio-economic, built environment, and mobility conditions associated with crime: a study of multiple cities. Sci Rep 10(1). https://doi.org/10.1038/s41598-020-70808-2

Newman O (1972) Defensible space. Macmillan New York
Peeters M, Vander Beken T (2017) The relation of cpted characteristics to the risk of residential burglary in and outside the city center of ghent. Appl Geogr 86:283–291. https://doi.org/10.1016/j.apgeog.2017.06.012
Qiu S, Psyllidis A, Bozzon A, Houben GJ (2019) Crowd-mapping urban objects from street-level imagery. In: The world wide web conference on-WWW '19, ACM Press. https://doi.org/10.1145/3308558.3313651
Reynald DM, Elffers H (2009) The future of newman's defensible space theory. European J Criminol 6(1):25–46. https://doi.org/10.1177/1477370808098103
Riemen-schneider H, Krispel U, Thaller W, Donoser M, Havemann S, Fellner D, Bischof H (2012) Irregular lattices for complex shape grammar facade parsing. In: 2012 IEEE conference on computer vision and pattern recognition, IEEE, pp 1640–1647
Rydin Y, Bleahu A, Davies M, Dávila JD, Friel S, De Grandis G, Groce N, Hallal PC, Hamilton I, Howden-Chapman P et al (2012) Shaping cities for health: complexity and the planning of urban environments in the 21st century. The Lancet 379(9831):2079–2108
Saville G, Cleveland G (2008) Second-generation cpted: The rise and fall of oppor- tunity theory. In: 21st century security and CPTED, Auerbach Publications, pp 93–104
Shach-Pinsly D (2019) Measuring security in the built environment: evaluating urban vulnerability in a human-scale urban form. Landsc Urban Plan 191:103412. https://doi.org/10.1016/j.landurbplan.2018.08.022
Sharifi Noorian S, Qiu S, Psyllidis A, Bozzon A, Houben GJ (2020) Detecting, classifying, and mapping retail storefronts using street-level imagery. In: Proceedings of the 2020 international conference on multimedia retrieval, ACM. https://doi.org/10.1145/3372278.3390706
Stafford M, Chandola T, Marmot M (2007) Association between fear of crime and mental health and physical functioning. Am J Public Health 97(11):2076–2081
Teboul O, Kokkinos I, Simon L, Koutsourakis P, Paragios N (2011) Shape grammar parsing via reinforcement learning. In: CVPR 2011, IEEE, pp 2273–2280
UN-Habitat (2020) United nations system-wide guidelines on safer cities and human settlements. Report
Won J, Lee C, Forjuoh SN, Ory MG (2016) Neighborhood safety factors associated with older adults' health-related outcomes: a systematic literature review. Soc Sci Med 165:177–186

Chapter 13
Automatic Evaluation of Street-Level Walkability Based on Computer Vision Techniques and Urban Big Data

A Case Study of Kowloon West, Hong Kong

Lu Huang, Takuya Oki, Sachio Muto, Hongjik Kim, Yoshiki Ogawa, and Yoshihide Sekimoto

Abstract The walkability of an urban environment is a critical aspect of urban design and planning, and has a direct impact on the quality of life for residents. Therefore, it is essential to conduct a systematic evaluation of the pedestrian environment to improve the walkability of a city. In recent years, there has been a growing emphasis on the application of automated evaluation methods, incorporating artificial intelligence and urban big data analysis. This study proposes a systematic walkability evaluation index with automated measurement capabilities, and corresponding measurement pipelines utilizing computer vision techniques as well as urban big data. To demonstrate the utility of the proposed index and measurement methods, this study conducts a systematic measurement of street-level walkability in the Kowloon West of Hong Kong as a case study.

Keywords Walkability evaluation · Street-level · Computer vision · Urban big data · Hong Kong

L. Huang (✉) · T. Oki
School of Environment and Society, Tokyo Institute of Technology, Tokyo, Japan
e-mail: huang.l.ac@m.titech.ac.jp

T. Oki
e-mail: oki.t.ab@m.titech.ac.jp

S. Muto · H. Kim
Center for Real Estate Innovation, The University of Tokyo, Tokyo, Japan
e-mail: s-muto@e.u-tokyo.ac.jp

H. Kim
e-mail: hj-kim@e.u-tokyo.ac.jp

Y. Ogawa · Y. Sekimoto
Center for Spatial Information Science, The University of Tokyo, Tokyo, Japan
e-mail: ogawa@csis.u-tokyo.ac.jp

Y. Sekimoto
e-mail: sekimoto@csis.u-tokyo.ac.jp

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_13

13.1 Introduction

13.1.1 Background

There is an increasing consensus that good walkability will create substantial benefits to the city and its citizens. In this context, the systematic evaluation of local pedestrian environments is a critical first step in upgrading urban walkability. Street-level walkability, encompassing fine-grained urban design features and street facility conditions, has captured the attention of urban designers and administrators in recent years, because it is simpler, faster, and cheaper to adjust built environments; thus, immediate interventions to increase street-level walkability are more practicable than those focused on macro and meso-scale characteristics (Millstein et al. 2013). Meanwhile, recent breakthroughs in urban big data and artificial intelligence (AI) technologies, particularly in computer vision (CV) techniques, have generated a plethora of options for rapid, large-scale evaluations of street-level pedestrian environments.

13.1.2 Related Studies

Over the last two decades, many studies have developed and validated walkability auditing tools focused on street-level factors. For example, Day et al. (2006) and Boarnet et al. (2006) investigated built environments that promote physical activity, ultimately resulting in the Irvine Minnesota Inventory (IMI), which consists of 162 detailed variables. Later, Millstein et al. (2013) developed the Microscale Audit of Pedestrian Streetscapes (MAPS), which aims to collect thorough data on the pedestrian environment. Focusing on conditions in Hong Kong, another study proposed two assessment checklists to measure street-level walkability, including one for citizens and one for professionals (Ng et al. 2016). Other street-level walkability indexes include the Analytic Audit Tool (Brownson et al. 2003), Walking Suitability Assessment Form (WSAF) (Emery et al. 2003), PIN3 Neighborhood Audit Instrument (Evenson et al. 2009), and neighborhood sidewalks assessment tool (NSAT) (Aghaabbasi et al. 2017). However, these indexes are designed for in-field studies or virtual auditing, and therefore require substantial human resources and time, which increases the difficulty of large-scale research.

Due to recent advances in geographic information system (GIS) technology and AI algorism, a wealth of opportunities exists for the automatic measurement of street-level walking environments. For example, Shammas and Escobar (2019), Yin (2017), and Limgomonvilas and Nimanong (2018) applied GIS technology and geodata to measure urban walkability. Yin and Wang (2016) applied Artificial Neural Network and Support Vector Machine to segment the sky pixels and studied the relationship of the visual enclosure with pedestrian numbers and Walk Scores. Zhou et al. (2019), Li et al. (2020) and Koo et al. (2022a) used semantic segmentation to segment streetscapes such as trees, obstacles, sky, sidewalk pavements, roads, buildings, and fences and used them to represent characters of micro walkability. Shortly after, Koo et al. (2022b) proposed applying instance segmentation with GIS to measure the presence status of micro streetscapes using the MAPS-min index, which includes 15 variables. To measure subjective aspects of streetscapes, an MIT research team proposed an urban perception analysis method based on crowdsourcing data and the convolutional neural network (CNN) (De Nadai et al. 2016). Oki and Kizawa (2022) expanded the capabilities of this method.

13.1.3 Current Gaps and Research Objectives

Although several street-level walkability indexes exist, and AI techniques have opened new avenues for evaluating street-level pedestrian environments, current research on automatic walkability evaluation is still preliminary. First, majority of existing walkability indexes cover rich variables for evaluating micro-walkability are formulated for manual evaluation, thus they cannot be easily inherited and implemented by automated evaluation studies. This is why majority of street-level automated evaluations often rely on a limited number of variables for which automated techniques have been demonstrated. Second, while the most popular technique for street-level automatic measurement, semantic segmentation, may capture the relative pixel ratio of background features, it is ineffective for the fine-grained analysis of foreground street facilities. Furthermore, there are still no practical measurement techniques for a number of important micro variables, including sidewalk width, sidewalk maintenance level, and pedestrian flow.

Given the gaps in the existing literature, this chapter aims to:

- Develop a new evaluation index that effectively balances the automatic measurability of street-level walkability variables with the comprehensive coverage of rich micro-level street characteristics.
- Propose new methodological pipelines to provide a more comprehensive and efficient measurement of street-level walkability variables.

13.2 Methodology

The overall methodological framework of this study is illustrated in Fig. 13.1. The framework begins with the selection of walkability dimensions and variables that possess automatic quantification potential to construct a comprehensive walkability index. Subsequently, data collection and measurement strategies were devised and implemented to collect and measure the chosen variables utilizing Kowloon West, Hong Kong as a case study. Finally, the CRITIC method was employed to determine the relative weights of each variable.

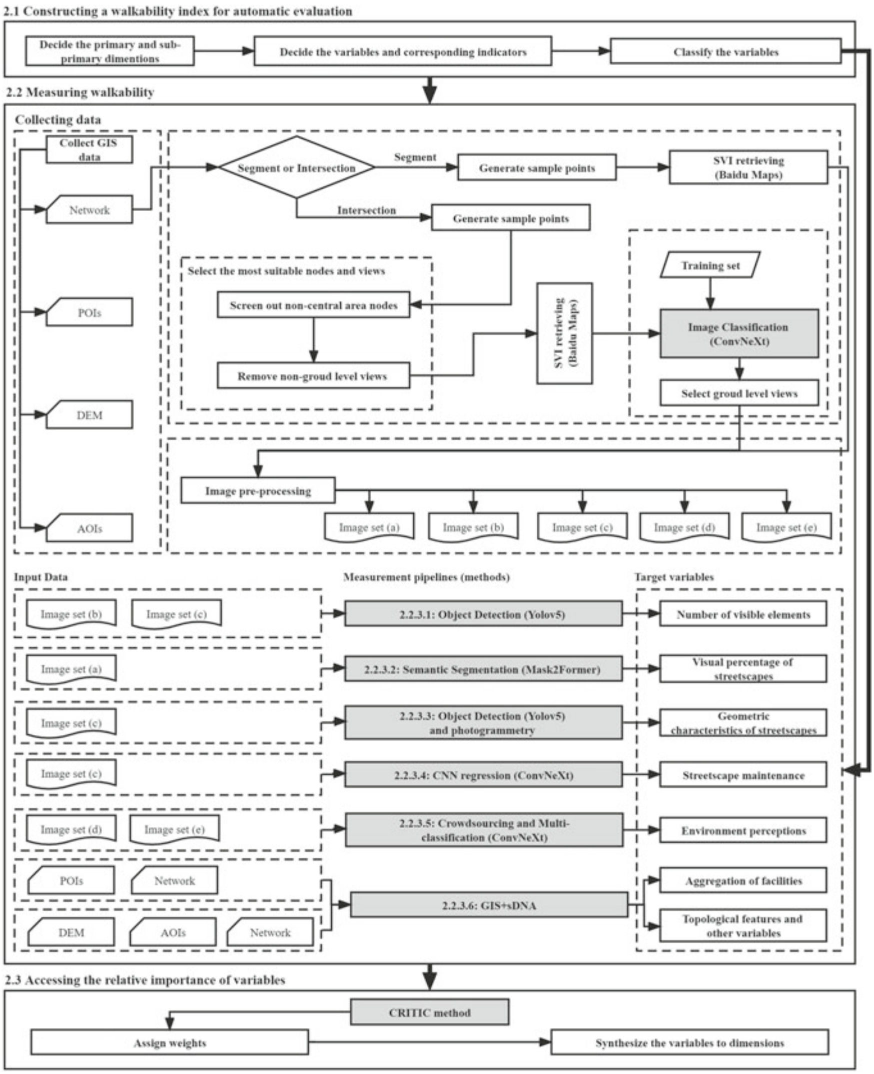


Fig. 13.1 Overall methodological framework

13.2.1 Constructing a Walkability Index for Automatic Evaluation

Because most street-level walkability indexes are not designed for automatic evaluation, this study created a more suitable version in two stages: (1) determine primary dimensions and subdimensions, and (2) select variables and the corresponding indicators. The new index features a multilevel structure that reflects specific and comprehensive characteristics, producing an easily discernible visualization of the evaluation results. It also distinguishes between street-segment and intersection spaces as well as objective and subjective walkability and screens the evaluation variables for each.

13.2.1.1 Selecting Primary Dimensions and Subdimensions

The primary dimensions and subdimensions were developed based on the Link-Place principle put forth by the Coalition (Coalition 2012) and the theoretical framework of protection, comfort, and pleasure proposed by Gehl Institute (2016) Based on these principles, we first constructed three primary dimensions for characterizing the general features of street-segment walkability, including whether they are safe to walk (A), comfortable to walk (B), and interesting to walk (C). As for street intersections, pedestrians typically express how well a particular arrangement accommodates their travel based on safety (Petritsch, et al. 2005). Thus, safe to walk (E) is prioritized when evaluating intersection walkability.

Subdimensions should reflect more specific evaluation aspects. This study selected 11 objective subdimensions for street-segment spaces, including accessibility (A-1), maintenance (A-2), protection against crime (A-3), protection from vehicles (A-4), protection from unpleasant weather (B-1), wind environment and ventilation (B-2), natural scenery (B-3), public realm amenities (B-4), destination (C-1), legibility & orientation (C-2), and active transportation (C-3). For intersection spaces, this study selected two objective subdimensions, including crossing safety (E-1) and crossing efficiency (E-2). Further, subjective walkability indicators included perceived safety (D-1 and F-1), perceived comfort (D-2), and perceived interest (D-3). This allowed us to compare subjective and objective walkability at an identical dimensional level.

13.2.1.2 Selecting Variables and Indicators

Since both primary dimensions and subdimensions are not quantifiable, it is necessary to obtain variables and corresponding indicators. Therefore, a small expert team composed of individuals conversant with CV or GIS technology was organized. First, experts employed existing indexes, primarily focusing on three widely recognized indexes: IMI (Boarnet et al. 2006), MAPS (Millstein et al. 2013), and CEx WALKScore (Ng et al. 2016), to screen for variables that could be quantified through CV or GIS techniques, resulting in 105 variables being identified. Next, they conducted a screening to eliminate variables with duplicate interpretations or which depended on overly complicated quantification methods, ultimately resulting in 50 (Tables 13.1 and 13.2).

In the process of deciding indicators of variables, we carefully evaluated and chose those indicators that have the capability to capture the street-level variations of the variables. For instance, previous research has established a correlation between the

Table 13.1 Selected dimensions, variables and indicators (segment)

	Primary dimension	Subdimension	Variables	Indicator of variables	Variable category	Positive/Negative
Objective	A. Safe to walk	A-1. Accessibility	A-1–1. Sidewalk width	Width of the sidewalk	C3	P
			A-1–2. Sidewalk steepness	Degree of slope	C3	N
			A-1–3. Curb cuts	Number of curb cuts	C1	P
			A-1–4. Connect to elevated walkway	Aggregation of the exit/entrance points	C6	P
			A-1–5. Connect to underground footway	Aggregation of the exit/entrance points	C6	P
		A-2. Maintenance	A-2–1. Sidewalk maintenance	Sidewalk maintenance score	C4	P
			A-2–2. Building façade maintenance	Building facade maintenance score	C4	P
		A-3.Protection against crime	A-3–1. Eyes of the shop owners	Length ratio of the street-side business	C3	P
			A-3–2. Eyes on the streets	Betweenness level of the street	C7	P
			A-3–3. CCTV camera	Number of CCTV cameras	C1	P
			A-3–4. Outdoor lighting	Number of lighting posts	C1	P
		A-4. Protection from vehicles	A-4–1. Sidewalk buffer	Length ratio of the fence	C3	P
			A-4–2. Painted crosswalk	Number of painted crosswalks	C1	P

(continued)

Table 13.1 (continued)

	Primary dimension	Subdimension	Variables	Indicator of variables	Variable category	Positive/Negative
			A-4–3. Traffic light	Number of traffic lights	C1	P
			A-4–4. Outdoor lighting	Number of lighting posts	C1	P
	B. Comfortable to walk	B-1. Protection from unpleasant weather	B-1–1. Shelter from and exposure to rain	Number of storefront awnings, and bus stop shelters	C1	P
			B-1–2. Shelter from and exposure to sunlight	Visual percentage of the sky	C2	N
		B-2. Wind environment and ventilation	B-2–1. Street parallel with the prevailing wind	Street parallel with the prevailing wind	C7	N
			B-2–2. Street FAR	Street FAR	C7	N
			B-2–3. Street overall width	Road grade of street network	C7	P
		B-3. Natural scenery	B-3–1. Greenery	Visual percentage of Greenery	C2	P
			B-3–2. Mountain scenery	Visual percentage of mountain scenery	C2	P
			B-3–3. Water scenery	Visual percentage of water scenery	C2	P
		B-4. Public realm amenities	B-4–1. Outdoor seating	Number of outdoor seatings	C1	P
			B-4–2. Post box	Number of post boxes	C1	P
			B-4–3. Trash-can	Number of post trash-can	C1	P

(continued)

Table 13.1 (continued)

	Primary dimension	Subdimension	Variables	Indicator of variables	Variable category	Positive/Negative
			B-4–4. Phone booth	Number of phone booth	C1	P
			B-4–5. Public toilet	Aggregation of public toilets	C6	P
	C. Interesting to walk	C-1. Destination	C-1–1. Park, open space, and tourism attraction	Aggregation of parks, open spaces, and tourism attractions	C6	P
			C-1–2. Public service facilities	Aggregation of public service facilities	C6	P
			C-1–3. Shopping or catering facilities	Aggregation of shopping or catering facilities	C6	P
		C-2. Legibility & orientation	C-2–1. Wayfinding system	Number of wayfinding systems	C1	P
			C-2–2. Banner	Number of banners	C1	P
		C-3. Active transportation	C-3–1. Biking facility	Presence of biking network	C7	P
			C-3–2. Bus stop	Aggregation of bus stops	C6	P
			C-3–3. Rail transit station	Aggregation of Rail transit station exit/entrance points	C6	P
Subjective	D. Environmental perception (segment)		D-1. Perceived safety	Perception score	C5	P
			D-2. Perceived comfort	Perception score	C5	P
			D-3. Perceived interest	Perception score	C5	P

Table 13.2 Selected dimensions, variables and indicators (intersection)

	Primary dimension	Subdimension	Variables	Indicator of variables	Variable category	Positive/Negative
Objective	E. Safe to walk	E-1. Crossing safety	E-1–1. Traffic light	Number of traffic lights	C1	P
			E-1–2. Yield sign/marking	Number of yield marking	C1	P
			E-1–3. Painted crosswalk	Number of painted crosswalks	C1	P
			E-1–4. Outdoor lighting	Number of lighting posts	C1	P
			E-1–5. Refuge island	Number of refuge islands	C1	P
			E-1–6. Curb cut	Number of curb cuts	C1	P
			E-1–7. Connect to elevated walkway	Aggregation of the exit/entrance points	C6	P
			E-1–8. Connect to underground footway	Aggregation of the exit/entrance points	C6	P
		E-2. Crossing efficiency	E-2–1. Crossing exposure	Road grade	C7	N
			E-2–2. Intersection complexity	Number of legs	C7	N
Subjective	F. Environmental perception (intersection)		F-1. Perceived safety	Perception score	C5	P

quantity of street amenities and walking intensity. Thus, quantifying the number of elements such as trash cans, seats, and lamp posts in a sampling area can provide a more nuanced understanding of the impact of built environment differences on walking behavior. However, it should be noted that quantifying the number of presences is not an appropriate measuring indicator for all variables, such as greenery, mountain scape, or water scape, as these variables' number cannot be effectively quantified. Therefore, we employed visual proportion as the indicator to assess these variables.

In order to overcome the difficulty of selecting direct indicators for certain variables, proxies were employed. One such example is the variable "eyes on the street" proposed by Jacobs (1961) as a crucial factor impacting street safety. However, it is challenging to directly measure the flow of people on the street. Therefore, we utilized "betweenness", a spatial syntax indicator, as a proxy. Betweenness is quantified as the frequency of traversal of street segment by the "shortest" path between any two other street segments within a specified analysis radius. This calculation reflects the potential for high throughput on that particular street segment. In addition to the natural surveillance provided by pedestrians on the streets, Jacobs (1961) also highlights the natural surveillance provided by owners and customers of street-side stores. To capture this, this study employed the "length ratio of storefronts" as a proxy indicator for the variable. For the variable "Shelter from and exposure to sunlight" was characterized by using the level of sky openness as a proxy to represent the shading opportunities provided by the street. For the variable "Crossing exposure", the road grade attribute of OpenStreetMap (OSM) road network was chosen as the proxy. The score assigned is inversely proportional to the grade of the road intersecting the intersection node.

To enhance the efficiency of the measurement strategy development process, we classified the variables based on their measurement characteristics, which resulted in the identification of seven categories of variables. The first category (C1) measures the number of foreground elements in the scene; The second category (C2) measures the visual percentage of background streetscapes; The third category (C3) measures the geometric characteristics of streetscapes; The fourth category (C4) measures the level of maintenance on streetscapes. The fifth category (C5) measures how people perceive the environment. The sixth category (C6) measures the aggregation of geo data. The final category (C7) measures topological characteristics, and other kinds of variables.

13.2.2 Measuring Walkability

This section presents an illustration of the methodology employed for data preparation and measurement using Kowloon West, Hong Kong as a case study site.

13.2.2.1 Study Site, Object, and Unit

The study site was Kowloon West, which is the western part of the Kowloon Peninsula in Hong Kong. It borders the sea to the west and south, with most of the land in the waterfront area being reclaimed. The main urban arteries run along the coastline. Topography to the north and northeast is relatively steep, as the site borders mountains. The central area is highly dense with buildings and streets.

Although Hong Kong contains one of the most developed multidimensional street networks in the world, the most serious conflicts between pedestrians and vehicles typically occur on streets that are located at ground-level and those which are equally shared by pedestrians and vehicles. These streets should clearly be prioritized for walkability improvements, and were therefore the main focus in this study. Regarding the study unit, we differentiated between street-segment and intersection spaces. For street-segment spaces, we set the segment (section) as the unit. Because some excessively long street segments may affect the accuracy of the results, we divided those measuring >100 m into several sections measuring ≤100 m, resulting in 4,786 selected segments. For pedestrian intersections, we set the junction node as the study unit, resulting in 1,229 selected nodes.

13.2.2.2 Collecting Street View Images and GIS Data

We obtained Street View images (SVI) (shooting period from 2017 to 2018) from Baidu Maps, which is a leading big data vendor in China. Different SVI collection procedures were used to acquire views of street-segment and intersection spaces. For street-segment spaces, we collected 13,983 panoramic images (4096*2048 pixels), with sample points generated at 20 m intervals along the street network. According to Gehl (1987), distances measuring 20 to 25 m may facilitate social encounters and fit the scale of 'place' in public open spaces.

When collecting SVI for intersections, utilizing coordinates for all OSM network junction nodes may not yield the desired results, as it could result in images of slip lanes rather than the view of central region of the intersection. Furthermore, it should be noted that the imagery retrieved through the API may not exclusively depict ground-level views, but may also include views from elevated highways or underground tunnels.

For the first problem, we identified whether network junction nodes were at the central intersection area by calculating the included angle between segments that converged at said nodes. Through examination of the nodes' characteristics, we determined that the majority of "suitable nodes" within our study site exhibited a connection to street segments with angles exceeding 30°. As a result, a threshold value was implemented, and a script was created to eliminate the nodes whose angle between segments is less than 30°. After cleaning, a total of 1299 nodes within the central region of the intersection remained. For the second problem, we collected excessive street views at each intersection node. Specifically, we created 10 m buffer radiuses using the intersection nodes after cleaning, then randomly generated several SVI collecting points therein. Thus, we gathered views using newly created collection points. Finally, we employed an image classification model, trained on two classes of images (ground-level view or non-ground-level view), to determine the presence of a ground-level view in the collected images.

GIS data, including street network, POIs, AOIs, and DEM were acquired from the OpenStreetMap and Hong Kong government's open data website: https://data.gov.hk/en/.

Table 13.3 Image dataset format

Dataset	Conversion parameters						Application
	Source area	Panorama to perspective	Heading (degree)	Field of View (degree)	Need crop or not	Image size (pixels)	
Image Set (a)	Segment	No	/	/	Yes	4096*1400	Measure visual percentage of streetscapes
Image Set (b)	Intersection	No	/	/	Yes	4096*1400	Measure the number of street elements (intersection)
Image Set (c)	Segment	Yes	360/180	90	No	1024*1024	Measure the number of street elements (segment)
Image Set (d)	Segment	Yes	270/90	90	Yes	1024*750	Measure the environment perceptions (segment)
Image Set (e)	Intersection	Yes	360/180 /270/90	90	Yes	1024*750	Measure the environment perceptions (intersection)

13.2.2.3 Pre-Processing SVIs and GIS Data

In alignment with the various types of image analysis tasks proposed in this study, different image strategies were utilized to pre-process the acquired SVIs. This resulted in the composition of five distinct image datasets (Table 13.3), each specifically tailored to address the corresponding analytical tasks.

13.2.2.4 Developing the Measurement Pipelines

To measure the seven categories of variables and indicators decided in Sect. 13.2.1.2, we developed seven corresponding measurement pipelines to quantify them.

Measuring the Number of Street Elements

In order to measure indicators of variables that signify the number of foreground street elements, we employed the widely utilized object detection algorithm YOLOv5

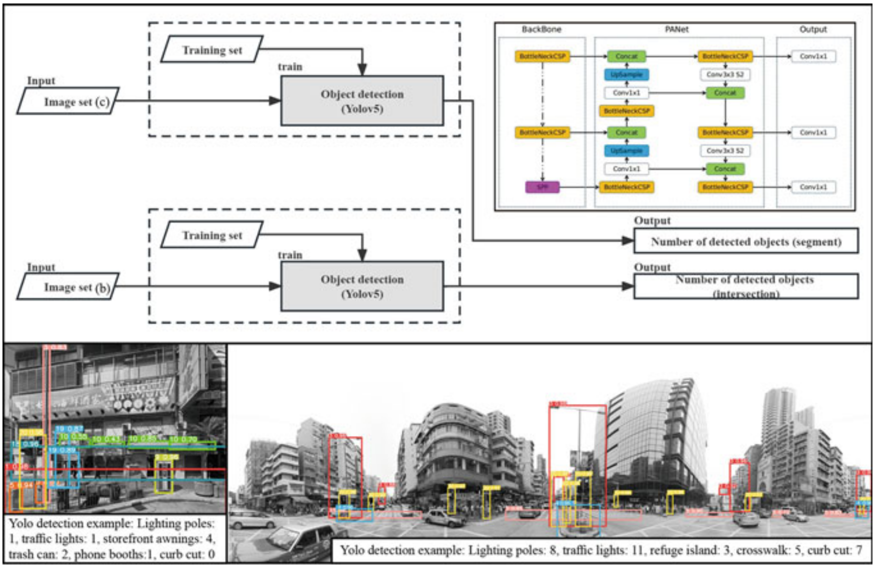


Fig. 13.2 Pipeline for measuring the number of streetscape elements

(Jocher et al. 2022) and utilized the perspective views captured from the street side (image set c) and panoramic images acquired at intersections (image set b) as input data for analysis (Fig. 13.2). Despite YOLOv5's availability of a substantial number of pre-trained weights, they are primarily trained on the COCO dataset (Lin et al. 2014) which contains a limited number of classes that are relevant to our research. Therefore, we selected 4000 street-side views as training images, and manually annotated them with a total of 13 new classes, including wayfinding systems, lamp posts, trash cans, post boxes, phone booths, storefront awnings, bus stop shelters, traffic lights, outdoor seating, curb cut, CCTV cameras, painted crosswalks, and banners. Similarly, 800 intersection panoramic images were selected and were annotated with 6 new classes, comprising of yield marking, curb cut, traffic lights, lamp poles, refuge islands, and painted crosswalks. After the annotation process was completed, we trained two new models using fine-tuning and pre-trained weights. The first model can be utilized to detect 13 classes of street elements in street segment areas, with a mean average precision (mAP) of 0.98 and 0.818 at a threshold of 0.5 and 0.95 respectively (precision = 0.962, recall = 0.959). The second model can be used to detect 6 classes of elements at street intersections, with a mAP of 0.986 and 0.808 at a threshold of 0.5 and 0.95 respectively (precision = 0.979, recall = 0.974).

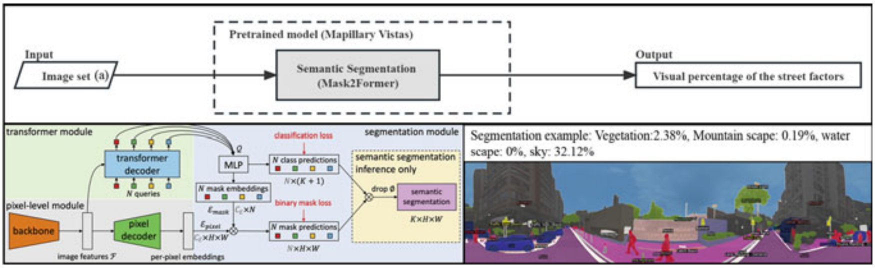


Fig. 13.3 Pipeline for measuring visual percentage of streetscape elements

Measuring the Visual Percentage of Streetscapes

In order to measure indicators of variables that signify the visual proportion of background streetscapes, we utilized processed panoramas (image set a) as the input for analysis.

To select an appropriate model, we employed a semantic segmentation algorithm, specifically, the pre-trained Mask2former model (Cheng et al. 2021) with ResNet50 as the backbone (Fig. 13.3). The Mask2former architecture is a novel approach that incorporates masked attention, which extracts localized features by constraining cross-attention within predicted mask regions. We selected the Mask2Former for this study due to its availability of a comprehensive library of pre-trained models (Cheng et al. 2022), which includes models trained on the Mapillary Vistas dataset that are capable of measuring visual percentage of vegetation, mountain scape, water scape and sky.

Measuring the Geometric Character of Street Elements

We proposed a methodological pipeline for quantifying the geometric characteristics of street elements based on the principles of similar triangles. This approach utilizes the relationship between the pixel sizes of SVIs and their corresponding distances in the real world to estimate the depth of the target subject (as illustrated in Fig. 13.4). In order to acquire information on the location of sidewalk pixels in images, we employed YOLOv5 and trained a model for detecting the range of the sidewalk pavement area. The model achieved a mAP of 0.99 at a threshold of 0.5 and 0.91 at a threshold of 0.95, with a precision and recall of 0.972 and 0.973, respectively. Subsequently, we used image set c as the input data and implement the estimation of the sidewalk width through the following steps. First, Eq. (13.1) was utilized to estimate the distance d from the camera's foot point to the edge of the road. In this equation, h represents the height of the vehicle camera, 512 represents half of the height of the images in pixels (since sidewalks are typically only visible in the lower half of the image), and n represents the number of vertical pixels of the road area in the image, which can be calculated using the parameters provided by the YOLO

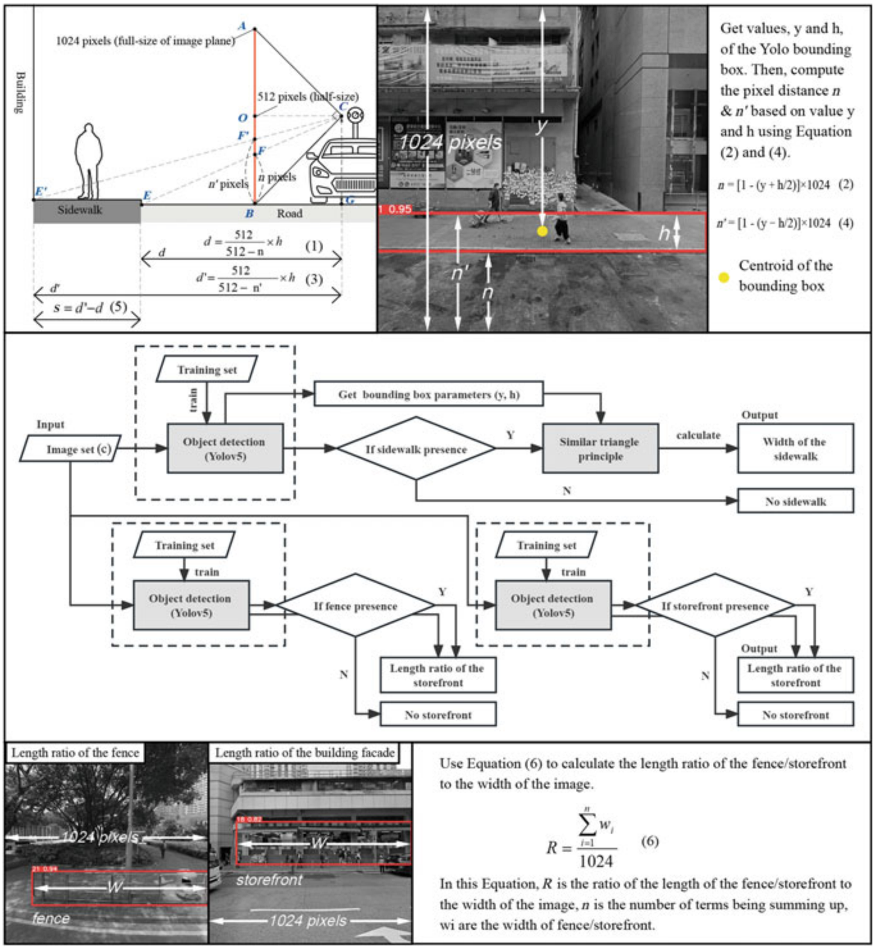


Fig. 13.4 Pipeline for measuring streetscape geometries

bounding box. These parameters include the ratio of the height (h) of the bounding box and the ratio of the distance from the centroid of the bounding box to the top edge of the image (y). Through the application of Eq. (13.2), the pixel distance of n can be determined. In the second step, Eq. (13.3) was utilized to estimate the distance d from the camera to the inside edge of the sidewalk. The value of n' can be calculated using Eq. (13.4). Finally, Eq. (13.5) was employed to measure the sidewalk width (w) by calculating the difference between *d'* and *d*.

The other two measured geometric indicators are the length ratio of street-side business (A-3-1) and fence (A-4-1). First, we performed manual annotation on the training set to label street-side storefronts and fences. Subsequently, two YOLOv5 models were trained using the annotated images. The width parameter of YOLO

Table 13.4 Evaluation criteria of building façade maintenance

Criteria	Description	Scoring scale
Aesthetical	Efflorescence; localized stains; uniform dirt; color change; runoff; graffiti; biological contamination; flatness deficiency; staining	1–5 (lowest—highest)
Adhesion loss and other facade defects	Cracking; detachments; peeling; erosion; blistering	
Physical disorder	Air conditioner external units are installed without any orders; Chaotic wiring arrangement; Stickers	

bounding box (w), was used to calculate the ratio of the sum of the length of detected bounding boxes to the width of the image frame (1024 pixels) using Eq. (13.6).

Measuring the Maintenance Levels of Street Elements

To evaluate the maintenance level-related variables, we employed a combination of CNN regression and YOLOv5 algorithms. The methodology was divided into three steps. The first step consisted of the development of scoring criteria for the maintenance level, and manual annotation of the training dataset based on these criteria. These criteria were designed in reference to existing literature and building façade/sidewalk maintenance specifications, as presented in Tables 13.4 and 13.5. To form the two training sets for the maintenance-level scoring model, a sample of 1000 images depicting building facades and 1000 images depicting sidewalks were selected, respectively, from image set c. Then, the two sets of images were scored manually based on criteria specific to building and sidewalk maintenance. The scoring process utilized a 5-point scale, where a score of 1 indicates poor maintenance and a score of 5 represents optimal maintenance. Finally, the final label score for each image was obtained by calculating the mean score of the three evaluation criteria. Subsequently, the annotated images were used as training data, and the corresponding scores were used as labels to train a CNN regression (maintenance-level scoring) model using ConVneXt (Liu et al. 2022) as the backbone, which is a novel architecture that employs advanced concepts from the Transformer model to enhance the performance of the traditional Res-Net50/200 network. The R-square values of the trained maintenance-level scoring models for building façade and sidewalk were 0.937 and 0.613 respectively, and the performance of the automatic scoring in the actual test was found to be consistent with expectations (Fig. 13.5).

The image set (c) was employed as the input data for the analysis of maintenance-related variables. However, it should be noted that the dataset utilized in the analysis may not always contain sidewalks or building façades. To address this issue, we employed the trained sidewalk detection model, in addition to new-trained model for

Table 13.5 Evaluation criteria of sidewalk maintenance

Criteria	Description	Scoring scale
Surface defects	Cracking; depressions; raveling; uneven surface	1–5 (lowest—highest)
Encroachment and obstruction	Overgrown plants encroaching on the sidewalk; Obvious piles of rubbish or litter on the pavement; presence of fixed obstructions such as tree pools, lamp poles or fire hydrants that have seriously obstruct walking activities; presence of temporary obstruction, such as temporary seating areas, garbage, and parking vehicles	
Curb defects	Cracking; detachments; erosion	

Fig. 13.5 Pipeline for measuring streetscape maintenance

building façade detection, to confirm the presence of sidewalks or building façades in the images. If sidewalks or building façades were present, the images were fed into the model for automatic scoring. If not, the images were assigned the average of the automatic scoring results during the visualization phase.

Measuring Environmental Perceptions

A methodology that combines crowdsourced data and advanced deep learning techniques was implemented to determine the key perceptual factors associated with walkability. We applied the data from a large-scale image pair comparison survey conducted by Ogawa et al. (2022). The survey was administered online via a mobile platform, where participants were randomly presented with two images on their mobile phones and asked to indicate which one they preferred based on 22 different questions, such as "Which street looks safer?". The study generated 14,950 pairs based on 1,500 randomly selected street images and received approximately 13,156,000 responses. The data was then processed using a CNN model with a ConVneXt as the backbone, converting the human subjective score extraction task from one logit value output to a multi-labeled classification. The model architecture is illustrated in Fig. 13.6. The original survey data was utilized to train the model in order to generate predictions for 22 dimensions of perception. Conversely, the evaluation prediction function, employed in this study, was limited to the utilization of only three dimensions, specifically safety, comfort, and interest. Analysis set d and e were utilized as the input data for this method pipeline.

Using GIS to Measure Aggregation-Related, Topology-Related and Other Variables

The utilization of CV techniques for the study of street-level built environments is highly beneficial, yet GIS methods remain advantageous for quantifying certain variables. In the analysis of aggregation-related variables, we employed GIS to calculate the average number of POI per meter of buffer zone for each street segment. Additionally, variables such as connectivity to elevated walkways (A-1-4) require the extraction of suspension points from the network of elevated walkways to represent overpass exits/entrances prior to aggregation. To analyze topology-related indicators such as Betweenness, we utilized Spatial Design Network Analysis (sDNA) (Cooper & Chiaradia, 2020), a leading software for spatial network analysis in GIS and CAD. The road network data underwent pre-processing and were subsequently input into the sDNA for calculation. Furthermore, we employed GIS to measure other variables, including crossing exposure (E-2-1) and intersection complexity (E-2-2) by calculating the average road grade and number of interacting segments at intersections. Additionally, we measured the performance of street-level ventilation by calculating the difference between the direction of the street segment and the

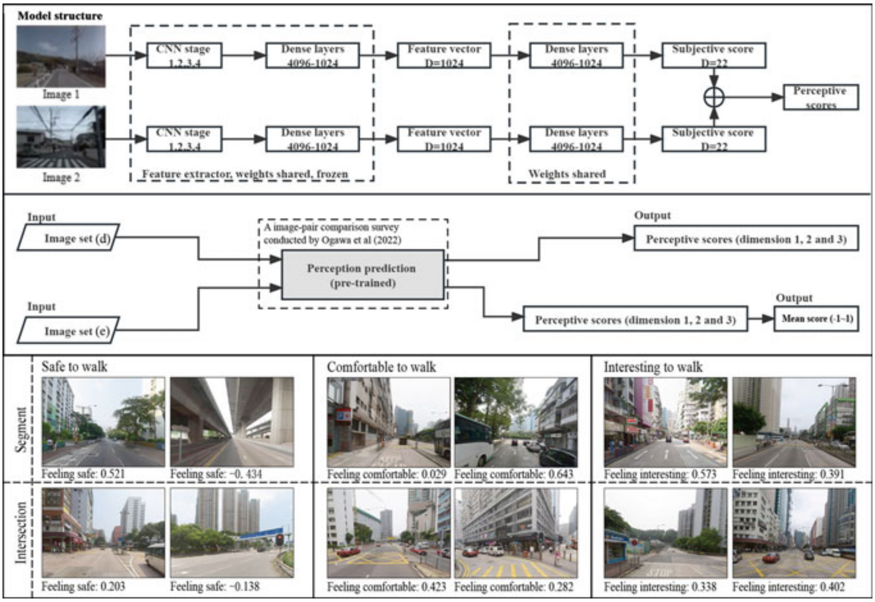


Fig. 13.6 Pipeline for measuring how people perceive the environment

direction of the prevailing wind in Kowloon West (B-2-1), with smaller differences indicating better performance.

13.2.3 Accessing the Relative Importance of Variables

Following the calculation of all variables, we utilized the CRITIC method to assess the relative importance of each variable. The method was proposed by Diakoulaki et al. (Determining objective weights in multiple criteria problems: The critic method, 1995), is an objective weighting method which aims to determine the objective weights of indicators based on the intensity of comparison and the conflict between the evaluation indicators. The calculation can be divided into four steps.

The initial step in the process is to identify the positive and negative attributes of the indicator and to subsequently conduct positive and negative normalization in accordance with these attributes, utilizing Eqs. (13.1) or (13.2).

$$x'_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \tag{13.1}$$

$$x'_{ij} = \frac{\max(x_j) - x_{ij}}{\max(x_j) - \min(x_j)} \tag{13.2}$$

The subsequent step is to compute the indicator volatility through the application of Eq. (13.3), where x_j represents the mean value of the data associated with each indicator.

$$S_j = \sqrt{\frac{\sum_{i=1}^{m}\left(x_{ij} - \overline{x}_j\right)^2}{n-1}} \tag{13.3}$$

The third step in the process is to determine the degree of conflict between indicators, which is calculated utilizing Eq. (13.4). In this equation, r_{ij} represents the correlation coefficient between the number of *i* indicator and the number of j indicator.

$$A_j = \sum_{i=1}^{n}\left(1 - r_{ij}\right) \tag{13.4}$$

Once the volatility and conflict have been quantified, the amount of information can be determined using Eq. (13.5). C_j represents the relative importance of the number of j evaluation index within the overall evaluation index system, with a larger value of Cj indicating a greater weight should be assigned to that index.

$$C_j = S_j \times A_j \tag{13.5}$$

Hence, the objective weight W_j of the number of j indicator can be calculated using Eq. (13.6).

$$W_j = \frac{C_j}{\sum_{j=1}^{n} C_j} \tag{13.6}$$

13.3 Experimental Results

13.3.1 Objective Walkability in Kowloon West

Utilizing the weights determined in Sect. 13.2.3, we were able to synthesize the 11 sub-dimensions that encapsulate the concept of walkability in the context of Kowloon West. As shown in Fig. 13.7, an obvious spatial pattern emerges on maps a, c, e, i, j, and k. High-value street segments were mainly clustered in central high-density blocks. This is because high-density environments continuously attract people and commercial resources. Driven by commercial interests, various facilities will automatically cluster in these areas. Increased economic values also stimulates the renewal of street facilities. Meanwhile, high-density urban forms provide more shelter from rain and reduce heat radiation, which greatly reduces the potential that pedestrians will be affected by bad weather.

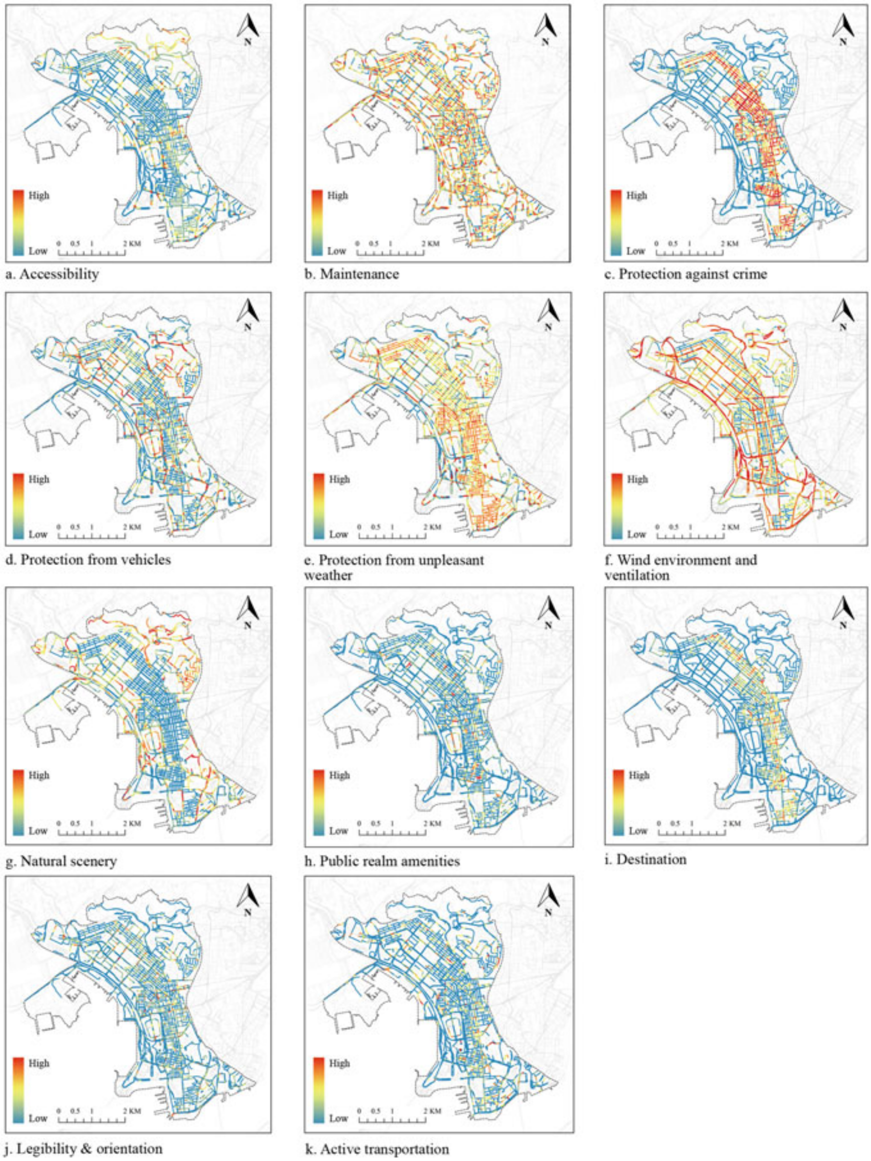
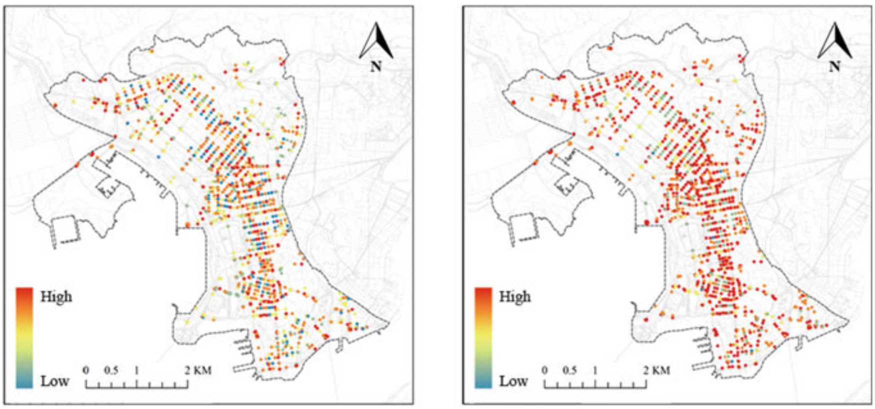


Fig. 13.7 Subdimensions of walkability in Kowloon West (segment)

We found opposite spatial distribution trends for street segments in maps f and g, the latter of which reflects the uneven spatial distribution of natural landscape in Kowloon West. High-density neighborhoods in the central area (e.g., Yau Ma Tei, Mong Kok, and Sham Shui Po) suffer from a severe lack of greenery and other natural landscapes. Further, ventilation problems in high-density blocks in Kowloon West must be taken seriously. As shown on map f, more segments with poor ventilation exist in high-density neighborhoods, which undoubtedly affects public health and summer walking experiences. As shown on maps k and j, respectively, ‘Active transportation’ and ‘Legibility & orientation’ are relatively evenly distributed throughout the study area, indicating that public transportation resources and wayfinding services are equitably accessible. While no clear spatial distribution pattern emerges on map b ‘maintenance’, it is roughly evident that segments with more low-maintenance features are mostly distributed in areas adjacent to main roads and neighborhoods that were built earlier. In regard to the intersection space, as depicted in Fig. 13.8l, m, the high-value points of map l, which pertains to crossing safety, are primarily situated on middle and high-grade streets. This suggests that the central region and intersections connected to higher-grade roads possess relatively superior pedestrian facilities. Map m, while displaying similar trends to map l, does not exhibit notably higher values for intersection crossing efficiency in the central region as compared to the surrounding areas.

To reflect walkability characteristics more comprehensively, we further aggregated the 11 subdimensions into four primary dimensions. Figure 13.9n–q shows the objective walkability results for Kowloon West at the primary dimensions, with visualizations of the spatial characteristics for safe to walk (n), comfortable to walk (o), and interesting to walk (p) in segment sections, as well as safe to walk (q) at intersections.



l. Crossing safety m. Crossing efficiency

Fig. 13.8 Subdimensions of walkability in Kowloon West (intersection)

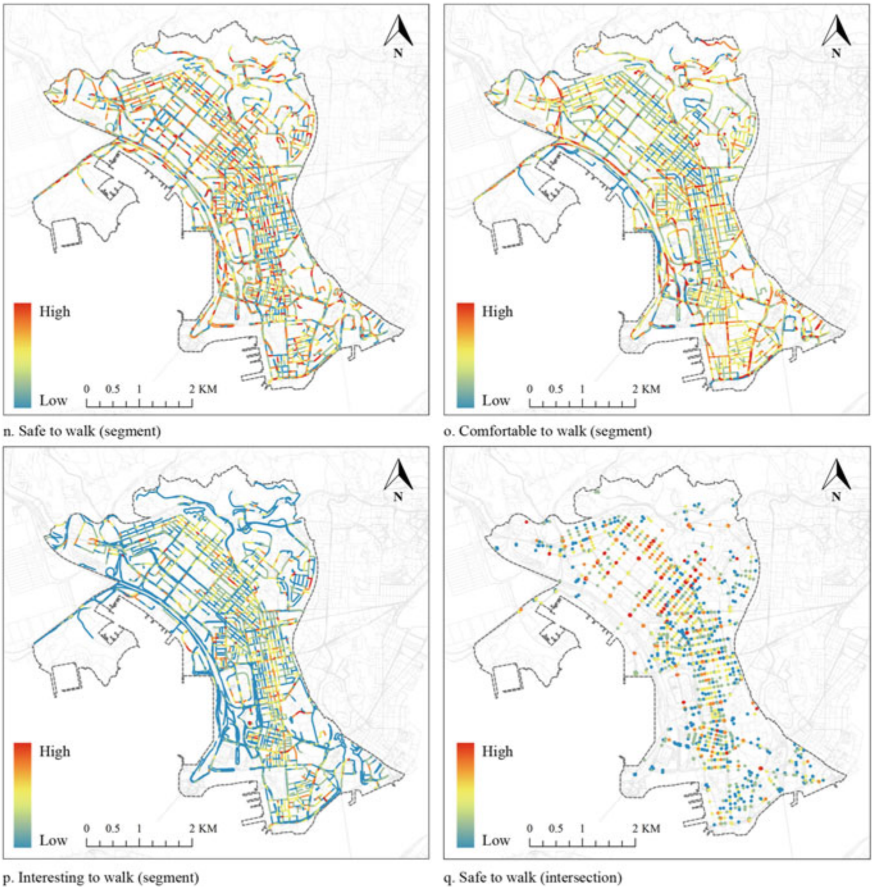


Fig. 13.9 Primary dimensions of walkability in Kowloon West (segment and intersection)

13.3.2 Subjective Walkability in Kowloon West

We used a trained perception prediction model to infer how pedestrians perceived the environments of street segments and intersections in Kowloon West. Figure 13.10 (r-u) shows the quantification results. For street segments, perceived comfort (s) and perceived interest (t) showed a spatial distribution similar to the measurement results for their corresponding objective dimensions, comfort to walk (o) and interesting to walk (p). However, we found a serious spatial mismatch between perceived safety (r and u) and its corresponding objective dimensions (n and q). For segment space, the dimension of the objective walking safety (safe to walk) revealed a lack of clustering of obvious high or low values. However, a distinct spatial pattern emerges when examining the perceived safety of streets in the Kowloon West. Street perceived as having high levels of safety tend to be located in peripheral areas, while those perceived as having low levels of safety are concentrated in central areas characterized in central areas characterized by earlier development and higher building density. In contrast, the spatial pattern of safety at intersections as measured by both subjective and objective methods was found to be opposite. This mismatch may have been influenced by the choice of variables. Moreover, the employed perception prediction method has its own limitations. First, it can only predict visual perception, and cannot reason on the full range of human perceptions of the built environment. Second, it cannot capture dynamic safety-related features (e.g., whether a street has a full-time stream of people). However, we cannot deny that people's perceptions of the built-environment are inherently biased to some extent (Gebel et al. 2009).

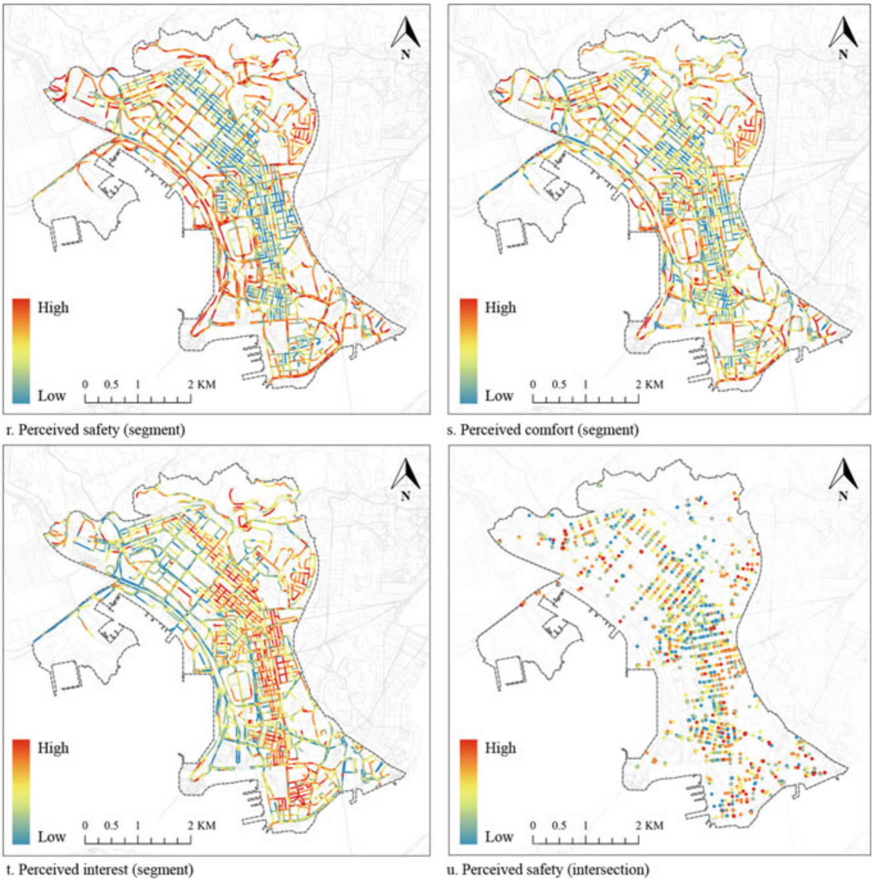


Fig. 13.10 Subjective dimensions of walkability in Kowloon West (segment and intersection)

13.4 Discussion

13.4.1 Can the Proposed Measurement Method Produce a Low-Cost, Fast, and Reliable Walkability Evaluation?

Applying urban big data, especially SVIs, can dramatically reduce the cost of conducting a street-level walkability evaluation. The Google Street View API service charges $7 USD per 1,000 images (Google Maps Platform 2022), while the Baidu Street View API service only charges ~$4.5 USD per 1,000 images. For the Kowloon West, the total street view acquisition cost is approximately $92 USD if using the Baidu Street View service. Because POI, street network data, and other geographical data are typically obtained for free or at minimal costs, they are considered negligible expenses. However, if the virtual audit approach is taken, assuming an average of 15 to 20 min for a proficient surveyor to audit each street segment/intersection, then it will take between 91,275 and 121,700 min (approximately1,521 to 2,028 h) to complete the audit tasks. If surveyors are paid the standard hourly wage in Hong Kong ($26.75 USD) (Salaryexplorer 2022), then the total cost for a virtual walkability audit of Kowloon West could reach $40,687 to $54,249 USD. If an in-field survey is adopted, then extra costs should be attached. Regarding evaluation speed and efficiency, virtual auditing and on-site surveys are substantially inferior to automatic evaluations based on CV and GIS. In terms of reliability, in-field surveys and virtual audits sometimes suffer from inconsistency when multiple surveyors assess streetscapes. However, the automatic walkability evaluation does not have this problem since the machine can fully follow the same instructions and criteria, thus producing more reliable results.

13.4.2 Do the Proposed Walkability Index and Its Methods Have High Applicability and Generalization Potential?

The applicability of the proposed index in different practical and research scenarios is at the beginning of index construction. It was built as a multilevel structure not only to reflect walking characteristics in different grain sizes, but also to broaden its applicability and target population. The higher the hierarchy of the index structure, the more comprehensive the variables and easier they are to grasp, thus rendering them ideal for activities such as civil participation planning, since the general public appreciates what ‘convenient to walk’, ‘safe to walk’, and ‘interesting to walk’ imply. The lower the hierarchy of the index structure, the more specialized knowledge is needed to interpret the dimensions/variables; those variables are more suitable for professional practitioners or researchers who guide specific urban design practices or conduct micro-level built-environment research.

Our index and measurement methods also have strong generalization potential due to the extensive coverage and consistent standard of applied urban big data. More than 140 nations/regions will have partial or complete coverage via street view images by 2022 (Google Maps Platform 2022). POI, street networks, and other urban big data have become widely available due to more developed electronic maps and matured crowdsourcing technologies. The proposed street-level walkability index and measurement techniques hold immense promise for transfer and application to other cities in this context. Nonetheless, there are still differences in the data sources and quality between cities/countries. Thus, any data gathering approach must be tailored to guarantee a seamless transfer.

13.4.3 What Are the Advantages of Applying Multiple AI Techniques in Walkability Evaluations?

Previous automatic walkability evaluations have generally adopted single analytic techniques. While such approaches are easy to implement, they also restrict the capacity to measure a variety of walkability variables. Further, the use of only one technique precludes fine-grained measurements. For example, it may be impossible to count the number of foreground street objects if solely using semantic segmentation to measure streetscape visibility. Meanwhile, it is difficult to determine the visual percentage of background stuff if solely using object detection. The use of a single analytic technique contradicts the point of objectively measuring street-level walkability, which needs as much detailed information of the built-environment as possible. In practice, walkability evaluations typically cover large scales, which may diminish the advantages of single techniques. As such, this study combined numerous CV techniques and GIS tools to capture as many detailed walkability characteristics as feasible.

13.5 Conclusion

This chapter develops a new index that strikes a balance between the richness of evaluation variables and the practicability of automated measurement. It also improves the methodology of walkability variable measurement. Using Hong Kong as a case study, we confirmed the practicality and dependability of the proposed index and its measurement techniques. Such research can give urban planners and policymakers viable references for creating walkability improvement plans, as well as provide potential factors for future research on the interaction between human behaviors and micro pedestrian environments.

However, this research also had some limitations. In terms of index construction, we omitted variables related to urban management (e.g., road speed limits and time schedules for vehicular street closures). While these can help determine walkability status, we excluded them due to a lack of relevant data. In terms of analyzing and explaining the evaluation results, we conducted an exploratory comparison between subjective and objective walkability, but a more in-depth examination of spatial parallels and mismatches is required.

Acknowledgements This work was supported by JST SPRING under Grant Number JPMJSP2106, and JSPS KAKENHI under Grant Number JP22K04490. We are grateful for the invaluable feedback provided by anonymous reviewers, which significantly enhanced the quality of this paper. Our sincere thanks to all involved parties.

References

Abdulla KM, Abdelmonem MG, Selim G (2017) Walkability in historic urban spaces: testing the safety and security in Martyrs' Square in Tripoli. Int J Archit Res Archnet-IJAR 11(3):163–177

Aghaabbasi M, Moeinaddini M, Shah MZ, Asadi-Shekari Z (2017) A new assessment model to evaluate the microscale sidewalk design factors at the neighbourhood level. J Transp Health 5:97–112

Al Shammas T, Escobar F (2019) Comfort and time-based walkability index design: a GIS-based proposal. Int J Environ Res Public Health 16(16):2850

Alves F, Cruz S, Rother S, Strunk T (2021) An application of the walkability index for elderly health—WIEH. The case of the UNESCO historic centre of Porto, Portugal. Sustainability 13(9):4869

Baidu Map Open Platform (2022) Static panorama. Retrieved 08 10, 2022, from https://lbsyun.baidu.com/index.php?title=viewstatic

Bartzokas-Tsiompras A, Photis YN (2020) Does neighborhood walkability affect ethnic diversity in Berlin? Insights from a spatial modeling approach. Europ J Geogr 11(1)

Belza B, Altpeter M, Smith ML, Ory MG (2017) The healthy aging research network: modeling collaboration for community impact. Am J Prev Med 52(3):228-S232

Boarnet MG, Day K, Alfonzo M, Forsyth A, Oakes M (2006) The Irvine-Minnesota inventory to measure built environments: reliability tests. Am J Prev Med 30(2):153–159

Brownson RC, Brennan Ramirez LK, Hoehner CM, Cook RA (2003) Analytic audit tool and checklist audit tool.

Cheng B, Misra I, Schwing AG, Kirillov A, Girdhar R (2022) Masked-attention mask transformer for universal image segmentation. CVPR

Cheng B, Schwing AG, Kirillov A (2021) Per-pixel classification is not all you need for semantic segmentation. NeurIPS

Coalition SA (2012) Streets for people: compendium for south Australian practice

Cooper CH, Chiaradia AJ (2020) sDNA: 3-d spatial network analysis for GIS, CAD. Command Line Python Softwarex 12:100525

Day K, Boarnet M, Alfonzo M, Forsyth A (2006) The Irvine-Minnesota inventory to measure built environments: development. Am J Prev Med 30(2):144–152

De Nadai M, Vieriu RL, Zen G, Dragicevic S, Naik N, Caraviello M, … Lepri B (2016) Are safer looking neighborhoods more lively? A multimodal investigation into urban life. In: Proceedings of the 24th ACM international conference on Multimedia, 1127–1135

Diakoulaki D, Mavrotas G, Papayannakis L (1995) Determining objective weights in multiple criteria problems: The critic method. Comput Oper Res 22(7):763–770

Dubey A, Naik N, Parikh D, Raskar R, Hidalgo CA (2016) Deep learning the city: quantifying urban perception at a global scale. European conference on computer vision, Springer(Cham), 196–212

Emery J, Crump C, Bors P (2003) Reliability and validity of two instruments designed to assess the walking and bicycling suitability of sidewalks and roads. Am J Health Promot 18(1):38–46

Evenson KR, Sotres-Alvarez D, Herring AH, Messer L, Laraia BA, Rodríguez DA (2009) Assessing urban and rural neighborhood characteristics using audit and GIS data: Derivation and reliability of constructs. Int J Behav Nutr Phys Act 6(1):44

Gebel K, Bauman A, Owen N (2009) Correlates of non-concordance between perceived and objective measures of walkability. Ann Behav Med 37(2):228–238

Gehl J (1987) Life between buildings. Van Nostrand Reinhold, New York

Gehl Institute (2016) PublicLifeDiversityToolkit_2.0

Google Maps Platform (2022) Google maps coverage. Retrieved 08 10, 2022, from https://developers.google.com/maps/coverage

Goolge Maps Platform (2022) Pricing that scales to fit your needs. Retrieved 08 10, 2022, from https://mapsplatform.google.com/pricing/

Hooper P, Knuiman M, Bull F, Jones E, Giles-Corti B (2015) Are we developing walkable suburbs through urban planning policy? Identifying the mix of design requirements to optimise walking outcomes from the 'Liveable Neighbourhoods' planning policy in Perth, Western Australia. Int J Behav Nutr Phys Act 12(1):1–11

Jacobs J (1961) The death and life of great american cities. Vintage, New York

Jacobs J (1992) The death and life of great american cities. Vintage, New York

Jocher G, Chaurasia A, Stoken A, Borovec J, NanoCode012, Kwon Y, …Minh M T. (2022) ultralytics/yolov5: v6.1—TensorRT. TensorFlow Edge TPU and OpenVINO Export and Inference, https://doi.org/10.5281/zenodo.6222936

Kim J, Park D, Suh Y, Jung D (2019) Development of sidewalk block pavement condition index (SBPCI) using analytical hierarchy process. Sustainability 11(24):7086

Koo BW, Guhathakurta S, Botchwey N (2022a) Development and validation of automated microscale walkability audit method. Health Place 73:102733

Koo BW, Guhathakurta S, Botchwey N (2022b) How are neighborhood and street-level walkability factors associated with walking behaviors? a big data approach using street view images. Environ Behav 54(1):211–241

Kostanjsek J, Lipar P (2007) Pedestrian crossings priority for pedestrian safety. 3rd Urban Street Symposium: Uptown, Downtown, or Small Town: Designing Urban Streets That WorkTransportation Research BoardInstitute of Transportation Engineers (ITE) US Access Board

Li Y, Yabuki N, Fukuda T, Zhang J (2020) A big data evaluation of urban street walkability using deep learning and environmental sensors-a case study around Osaka University Suita campus. eCAADe, 2, 319–328

Limgomonvilas T, Nimanong R (2018) Multi-criteria analysis and network analysis for walkability score in Amphoe Muang, Nonthaburi, Thailand. J Adv Res Soc Sci Humanities 3(4):125–135

Lin T-Y, Maire M, Belongie S, Bourdev L, Girshick R, Hays J, … Dollár P (2014) Microsoft COCO: common objects in context. Computer Vision–ECCV 2014: 13th European Conference, Springer International Publishing, pp 740–755

Liu Z, Mao H, Wu C-Y, Feichtenhofer C, Darrell T, Xie S (2022) A convnet for the 2020s. Proceedings of the IEEE/cvf conference on computer vision and pattern recognition

Millstein AR, Cain KL, Sallis JF, Conway TL, Geremia C, Frank LD, Saelens BE (2013) Development, scoring, and reliability of the Microscale Audit of Pedestrian Streetscapes (MAPS). BMC Public Health 13(1:1–15

Ng S, Lai C, Liao P, Lao M, Lau W (2016) Measuring and improving walkability in Hong Kong

Ogawa Y, Oki T, Zhao C, Sekimoto Y, Shimizu C (2022) Evelopment of a model for evaluating subjective impressions of streetscapes using omnidirectional street image big data. Okinawa: The 31st annual conference of geographical information systems association

Oki T, Kizawa S (2022) Model for evaluating impression of streets in residential areas based on image big data and a large questionnaire survey. J Archit Planning 87(800):2102–2113

Petritsch TA, Landis BW, McLeod PS, Huang HF, Challa S, Guttenplan M (2005) Level-of-service model for pedestrians at signalized intersections. Transp Res Rec 1939(1):54–62

Salaryexplorer (2022) Salaryexplorer. Retrieved 08 10, 2022, from http://www.salaryexplorer.com/salary-survey.php?loc=97&loctype=1#:~:text=The%20average%20hourly%20wage%20%28pay%20per%20hour%29%20in,into%20two%20categories%3A%20salaried%20jobs%20and%20hourly%20jobs

Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision, 618–626

Tchinda PE, Kim SN (2020) The paradox of "eyes on the street": pedestrian density and fear of crime in Yaoundé. Cameroon. Sustainability 12(13):5300

Transport for London (2018) Walking action plan: Making London the world's most walkable city

Wang H, Silva EA, Liu L (2022) Large-scale evaluation of the urban street view with deep learning method. Machine learning and the city: applications in architecture and urban design, 495–499

Wilson JQ, Kelling GL (1982) Broken windows. Atlantic Monthly 249(3):29–38

Yin L (2017) Street level urban design qualities for walkability: combining 2D and 3D GIS measures. Comput Environ Urban Syst 64:288–296

Yin L, Wang Z (2016) Measuring visual enclosure for street walkability: using machine learning algorithms and Google Street View imagery. Appl Geogr 76:147–153

Zhou H, He S, Cai Y, Wang M, Su S (2019) Social inequalities in neighborhood visual walkability: Using street view imagery and deep learning technologies to facilitate healthy city planning. Sustain Cities Soc 50:101605

Chapter 14
Promoting Sustainable Travel Through a Web-Based Tourism Support System

Yudai Kato and Kayoko Yamamoto

Abstract In Japan, distributed travel is recently promoted in order to prevent both the problems associated with overtourism and the spread of the COVID-19 pandemic in urban tourist destinations. The present study developed a tourism support system by integrating web-geographic information system (Web-GIS), recommendation system and social network services (SNS). The system has two unique key functions (the functions of tourism congestion display and tourist attraction recommendation) in order to promote distributed travel during the sightseeing planning stage. The system was operated for six weeks targeting Kamakura City in Kanagawa Prefecture, Japan. The evaluation results for the system performance revealed that the function of tourism congestion display can promote tourism that takes congestion periods and areas into consideration. It also showed that the function of tourist attraction recommendation can provide users with novel tourist attraction recommendations and achieve high levels of intent to visit recommended tourist attractions.

Keywords Tourism support system · Web-geographic information system (Web-GIS) · Recommendation system · Social network services (SNS) · Overtourism · Distributed travel

14.1 Introduction

Prior to the COVID-19 pandemic, the number of foreign visitors to Japan was rising year-by-year, and urban tourist destinations faced various problems as a result of overtourism. In Japan, when the pandemic comes to an end and the tourism industry recovers, the number of tourists is forecast to begin climbing once again, and the problems associated with overtourism are expected to return. According to Bongki

Y. Kato · K. Yamamoto (✉)
Graduate School of Informatics and Engineering, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu City, Tokyo 182-8585, Japan
e-mail: kayoko.yamamoto@uec.ac.jp

Y. Kato
e-mail: k2030031@edu.cc.uec.ac.jp

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_14

(2018) and Choi (2020), the fundamental cause of overtourism is the intake of large tourist numbers due to tourism promotion. To help address the problem associated with overtourism, the Japan Tourism Agency (2019) has recommended providing tourists with congestion information in a visible form. The United Nations World Tourism Organization (UNWTO) (2016) has recommended promoting the dispersion of visitors inside and outside urban areas and visits across different times of day. Based on these recommendations, the promotion of dispersion to discourage tourist traffic would be effective in helping prevent the problems associated with overtourism.

One way of promoting greater dispersion of tourists is distributed travel. In 2020, the Japan Tourism Agency launched a distributed travel promotional campaign to recommend that people go on trips while avoiding congestion by traveling at different times and to different destinations. This campaign recommends avoiding the three Cs of closed spaces, crowded spaces and close-contact settings in order to prevent the spread of the COVID-19 pandemic. Additionally, it is also effective in helping prevent overtourism that is caused by a large influx of visitors. Therefore, it is can be said that promoting distributed travel will be an important method of preventing overtourism that will occur following the COVID-19 pandemic.

Taking the above into consideration, the present study aims to develop a tourism support system that has two unique key functions of tourism congestion display and tourist attraction recommendation in order to promote distributed travel during the sightseeing planning stage. The system integrates web-geographic information system (Web-GIS), recommendation system and social network services (SNS). During the operation, the system has various users, both inside and outside the operation target area, who use the system and answer a web questionnaire survey. Based on the results of the questionnaire survey for users, the system evaluation is conducted.

Kamakura City in Kanagawa Prefecture, which is one of popular urban tourist destinations in Japan, is selected as the operation target area. The first reason for this is that the problems associated with the overtourism such as traffic congestion and tourists' bad manners occur, according to Kamakura City (2016). The second reason is that many tourists tend to visit this city in specific seasons. By providing users with visible information concerning crowded seasons and places, it is expected that the system can encourage them to consider appropriate timing and locations, when they plan to go sightseeing. The third reason is that tourists tend to visit specific popular tourist attractions. By providing users with information concerning tourist attractions that are novel and tailored to their preferences and tourism purposes, it is expected that the system can guide them to disperse.

14.2 Related Work

As mentioned in Sect. 14.1, the present study develops a unique system by integrating multiple systems such as Web-GIS, recommendation system and SNS. Therefore, considering the system characteristics, the present study is closely related to three study fields, namely, (1) studies related to dispersing tourists by adopting information processing techniques, (2) studies related to GIS-based tourism support systems, and (3) studies related to tourism support systems that adopt recommendation systems. The following introduces the major preceding studies of recent years in all parts of world in the above three study fields and demonstrates the originality of the present study in comparison.

In (1) studies related to dispersing tourists by adopting information processing techniques, as distributed travel was proposed, and it is recently promoted and discussed in Japan, this theme has been seldom researched in other countries. Kawai et al. (2018) and Himono et al. (2020) proposed systems that visually display the amount of congestion by using location information of tweets and users' behavior histories. Ma (2019) introduced analysis methods and utilization technologies of user-generated content (UGC) that are adopted to individualize and disperse tourism. Totsuka et al. (2021) conducted agent simulation that aimed to disperse visitors at tourist destinations during times of heavy congestion.

Regarding (2) studies related to GIS-based tourism support systems, Chang et al. (2011), Singh et al. (2011), Masron et al. (2015, 2016), Bali et al. (2016), Itou et al. (2018) and Koga et al. (2021) developed tourism support systems that display various information concerning tourist attractions and routes on the digital maps of Web-GIS. Kurata (2012), Kurata et al. (2015), Zhuang et al. (2015), García-Palomares et al. (2015), Sugimoto (2018), Kaufmann et al. (2019) and Hirota et al. (2019) created tourism potential maps that displayed areas with high tourism potential on the digital maps of Web-GIS, quantifying tourist destination highlights by using data extracted from photo-sharing sites and social media.

Regarding (3) studies related to tourism support systems that adopt recommendation systems, Cao et al. (2010), Kurashima et al. (2010), Majid et al. (2013, 2015), Hidaka et al. (2020) and Sumitomo et al. (2020) developed tourist attraction recommendation systems that utilized word of mouth information and images concerning tourism on the Internet. Noguera et al. (2012), Ikeda et al. (2014), Mizutani and Yamamoto (2021), An and Moon (2019), Kato et al. (2020), Esmaeili et al. (2020) and Abe et al. (2021) developed tourism support systems that recommended tourist attractions based on the degree of similarity between the feature vectors of users and tourist attractions.

A tourism support system developed in the present study has a function that uses GIS to visually display the tourism congestion in each district of the operation target area for each month on a digital map, together with a function that recommends users tourist attractions that are novel and tailored to their preferences and tourism purposes adopting recommendation system. Therefore, comparing the results of the preceding studies in the three study fields as listed above, the originality of the present study is to develop the system that provides users with detailed information concerning congestion periods and areas, recommends them appropriate tourist attractions, and promotes distributed travel during the sightseeing planning stage.

14.3 System Design

14.3.1 System Overview

As shown in Fig. 14.1, the system of the present study is developed by integrating Web-GIS, recommendation system and SNS. The purpose of the system is to promote distributed travel during the sightseeing planning stage. It achieves this by using the function of tourism congestion display adopting Web-GIS and the function of tourist attraction recommendation adopting recommendation system. The function of tourism congestion display is to display tourism congestion in each district of the operation target area for each month on a digital map. This provides users with visual information concerning when and where congestion occurs, which helps encourage sightseeing planning that takes appropriate timing and locations into consideration. The function of tourist attraction recommendation recommends users tourist attractions that are novel and tailored to their preferences and tourism purposes. This helps reduce the likelihood of the creation of sightseeing plans that only cover popular tourist attractions. Additionally, by integrating SNS and Web-GIS, it is possible for users to submit tourist attraction information, and display and share it on a digital map.

The system uses a web server, a database server and a GIS server. Heroku, which is a PaaS provided by the Salesforce, Inc., is adopted for both the web server and the database server. ArcGIS Online provided by the Environmental Systems Research Institute, Inc. (ESRI) is adopted for the GIS server. The main languages used to develop the system are Python, HTML and JavaScript. Additionally, the system is envisioned for use by both PCs and mobile information terminals. There are no differences in the functions of the system between devices, and users can use the same functions from any device.

14.3.2 Design of Each System

14.3.2.1 Web-GIS

The system adopts ArcGIS API for JavaScript provided by the ESRI to display the locations of tourist attractions, eateries and public transportation (railway stations and bus stops), the administrative boundaries of districts and cities, and the tourism congestion that will be inferred in Sect. 14.4.2 on digital maps, because it requires no software installation and can be used by accessing a website. Additionally, ArcGIS Pro provided by the ESRI is adopted to create layers in advance, and these are uploaded to the GIS server and displayed on a digital map. Five layers, which respectively contain information concerning tourist attractions, eateries, public transportation, the administrative boundaries and the tourism congestion, are created. When

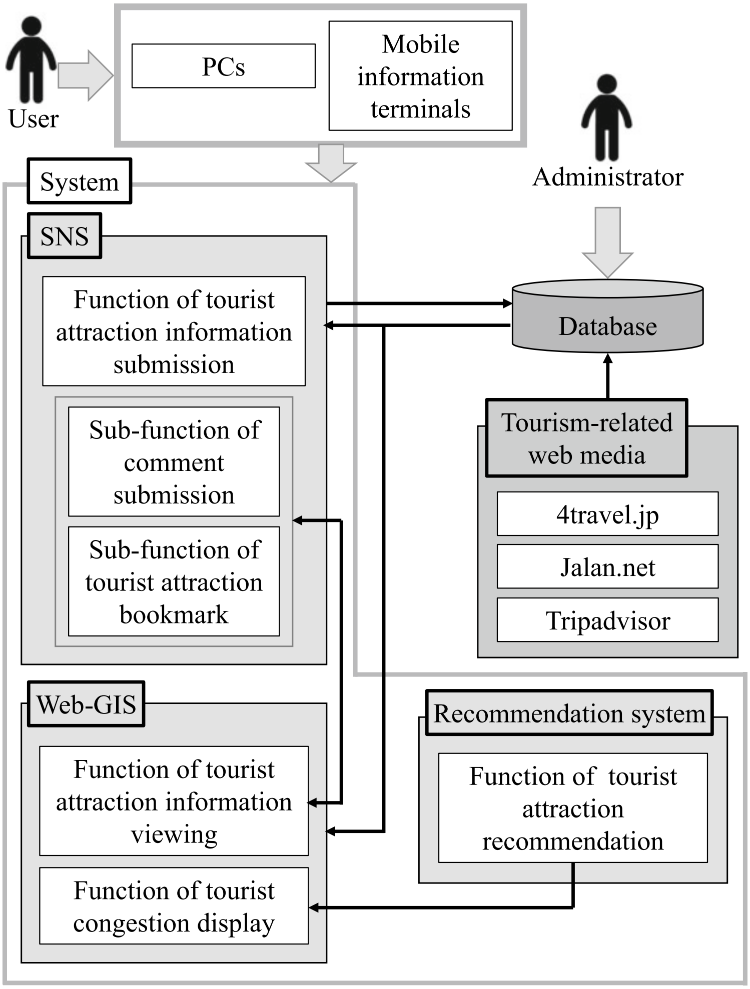


Fig. 14.1 System design of a tourism support system to promote distributed travel adopting GIS and recommendation system

displaying the above layers, Leaflet, which is an open source JavaScript library for map, is adopted to prevent the loading of the layer from slowing down processing.

14.3.2.2 Recommendation System

According to Kamishima (2008) and Jannach et al. (2010), there are three methods, including cooperative recommendation, content-based recommendation and knowledge-based recommendation, that can make recommendations from massive amounts of data to match users' preferences. Knowledge-based recommendation is selected for the recommendation system adopted in the system. The reason for this is that the method explicitly asks users about their preferences before making recommendations, and it has high potential for recommending tourist attractions where users will wish to visit. According to Kamishima (2008), utilizing knowledge-based recommendation also solves the cold start problem. This problem occurs when a lack of sufficient past information prevents a system from making appropriate recommendations for users. With the method, user information is developed by explicitly asking users about their preferences in advance. Using this information, appropriate recommendations can be made for users, even if there is no past information of them.

The system is envisioned for use by people of wide generations both inside and outside the operation target area. Therefore, it would be ideal to design the system to make recommendations simply by inputting information concerning users' conditions required for tourist attractions, using an explicit and intuitive question-and-answer entry format that requires no prior knowledge. In this format, users indicate the degree of importance placed on each condition required for tourist attractions, using a ten-point scale from 1 to 10. The values they entered are used to create a user profile, and these are used to define the feature vectors of users. For tourist attractions, feature values are calculated in advance for the same conditions that are required by users and accumulated in the database of the system. When making a recommendation, for tourist attractions, the feature values are extracted from the database of the system, and the feature vectors are created on the basis of them. The calculation of the feature values of tourist attractions is explained in Sect. 14.4.3. Equation (14.1) is applied to calculate the degree of similarity between the feature vectors of a user and a tourist attraction. The tourist attractions with high degree of similarity are recommended to users.

$$Sim_i = \frac{\sum_{j=1}^{n} \frac{I_j}{2} \times F_{ij}}{\sqrt{\sum_{j=1}^{n} \left(\frac{I_j}{2}\right)^2} \times \sqrt{\sum_{j=1}^{n} (F_{ij})^2}} \tag{14.1}$$

Sim_i: Degree of similarity between the degree of importance placed by user and the feature value of tourist attraction i

I_j: Degree of importance placed by user on condition j

F_{ij}: Feature value of tourist attraction i for condition j

i: Tourist attraction number

j: Condition number.

14.3.2.3 SNS

For the system, a unique SNS is designed. The main functions of the SNS are viewing and submitting information including images and comments, and bookmarking favorite tourist attractions of users. Therefore, on my page (each user's private page), they can check and delete submitted information and confirm bookmarked tourist attractions by them. Additionally, they can check and modify their personal information (their account names, genders, age brackets, and email addresses and passwords).

14.4 Database Creation of the System

14.4.1 Data Used in the System

Table 14.1 shows the data used in the system. Data concerning 107 tourist attractions, 167 eateries and 253 public transportations (24 railway stations and 229 bus stops) and the administrative boundaries of districts and cities in the operation target area were used and respectively processed in four layers adopting ArcGIS pro. Additionally, to implement the two unique key functions (the functions of tourism congestion display and tourist attraction recommendation), as shown in Table 14.1, reviews collected from tourism-related web media are utilized. The contents of the reviews are used to infer congestion periods and areas in Sect. 14.4.2, and these results are utilized in the function of tourism congestion display. Words extracted from the reviews are also utilized to calculate the feature values of tourist attractions in Sect. 14.4.3, and these results are utilized in the function of tourist attraction recommendation.

Table 14.1 Data used in the system

Data	Format	Related section	Sources
Reviews concerning tourist attractions	Text data	Sections 14.4.2 and 14.4.3	4travel.jp provided by the Kakaku.com, Inc., Jalan.net provided by the Recruit Co. Ltd., and Tripadvisor provided by the Tripadvisor LLC
Tourist attractions	Point data	Sections 14.4.3 and 14.4.4	
Eateries	Point data	Section 14.4.4	Kamakura City website
Public transportations	Point data	Section 14.4.4	Digital national land information provided by the Geospatial Information Authority of Japan
Administrative boundaries of districts and cities	Line data	Sections 14.4.2 and 14.4.4	

14.4.2 Inference of Tourism Congestion

Though tourist location information must be collected in order to infer tourism congestion, it is personal information that is difficult to obtain. In the present study, visit date information was extracted from reviews submitted to tourism-related web media that were introduced in Sect. 14.4.1, and then combined with related tourist attraction location information to infer the tourism congestion in each district of the operation target area for each month. The process to infer the tourism congestion in the present study is shown below. In steps (3)–(5), ArcGIS pro was adopted to create layers.

Step (1) 30,356 reviews concerning tourist attractions in the operation target area were collected from tourism-related web media.

Step (2) Visit date information was extracted from the reviews collected in step (1), and processed it in order to create visit number data concerning each tourist attraction for each month.

Step (3) A layer was temporarily created by combining the visit number data concerning each tourist attraction for each month created in step (2) and the related tourist attraction location information.

Step (4) Administrative boundary layer, which was introduced in Sects. 14.3.2.1 and 14.4.1, was overlaid with the layer created in step (3) to create a new one that contains the number of visits to each district for each month.

Step (5) For each district, Eq. (14.2) was used to determine the percentage of the number of visits for each month in the total number of annual visits, referring to the data included in the layer created in step (4). The value was used as the tourism congestion in each district for each month. In the present study, the tourism congestion in 44 districts where one or more tourist attractions are located was inferred. The reason for this is that tourist attractions concentrate and great many tourists frequently stay in the districts that are heavily congested. Additionally, two or more tourist attractions are located in most of the districts. Natural classification was used to divide the tourism congestion into six classes, and the result was used to create a new layer, which was uploaded to the GIS server.

$$C_{ij} = \frac{n_{ij}}{N_i} \tag{14.2}$$

C_{ij}: Tourism congestion in district i for month $.j$
N_i: Total number of annual visits to district i
n_{ij}: Number of visits to district i for month j
i : District number
j: Month number

14.4.3 Calculation of the Feature Values of Tourist Attractions

Reviews submitted to tourism-related web media were used to calculate the feature values of tourist attractions. Using data of reviews that were written by tourists, made it possible to extract information concerning conditions required for tourist attraction that are of high importance to them. Therefore, in the present study, words within reviews were extracted and utilized to calculate the feature values of tourist attractions. The process to calculate the feature values of tourist attractions in the present study is shown below.

Step (1) Words including noun, adjective, adverb and verb were extracted from 30,356 reviews collected in step (1) of Sect. 14.4.2, by conducting morphological analysis adopting MeCab, which is an open source engine.

Step (2) Expressions that indicate the features of tourist attractions, and tourists' preferences and tourism purposes were defined as feature categories, referring to the words extracted in step (1). The feature categories were congestion (high), congestion (low), access (difficult), access (easy), nature tourism, and historical and cultural tourism.

Step (3) For each tourist attraction, the words extracted in step (1) that fell within the feature categories defined in step (2) were picked out. These were respectively classified into each feature category by five persons.

Step (4) For each tourist attraction, the number of appearance frequency of the words classified into each feature category in step (3) was calculated.

Step (5) For each tourist attraction, Eq. (14.3) was used to determine the percentage of the number of appearance frequency of the words classified into each feature category calculated in step (4) in the number of reviews, and then normalize the result so that it could be more easily used by the system. The values that were obtained through this step were the feature values of tourist attractions. The maximum feature values of tourist attractions were set to 5. Any feature values that exceeded 5 were changed to 5.

$$F_{ij} = \frac{n_{ij}}{N_i} \times 5 \tag{14.3}$$

F_{ij}: Feature value of tourist attraction i for feature category $.j$
N_i: Number of reviews concerning tourist attraction i
n_{ij}: Number of appearance frequency of the words extracted from reviews concerning tourist attraction i for feature category j
i: Tourist attraction number
j: Feature category number

14.4.4 Database Creation

The feature values calculated in Sect. 14.4.3 were added to data concerning tourist attractions. Data concerning tourist attractions, eateries, public transportation, the administrative boundaries of districts and cities, and the tourism congestion that was inferred in Sect. 14.4.2 were accumulated into the database of the system. As mentioned in Sect. 14.3.2.1, all of the above data were respectively processed into five layers adopting ArcGIS pro. These layers were uploaded to the GIS server.

14.5 System Development

14.5.1 System Frontend

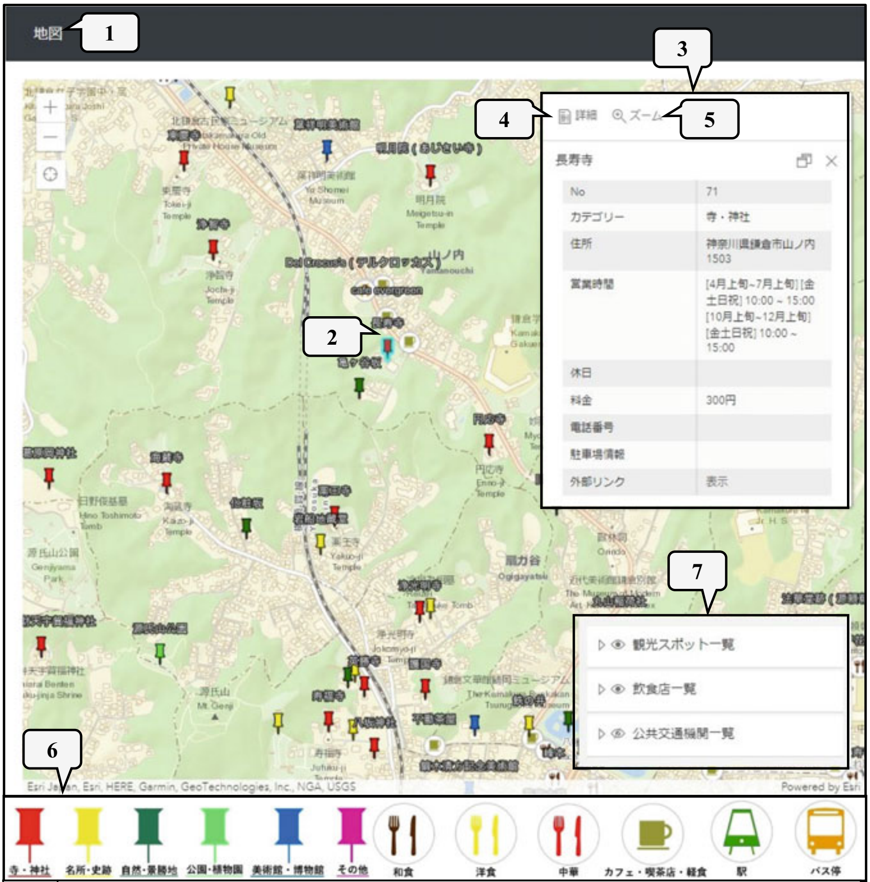
As the system of the present study was developed by integrating multiple systems, the following four unique functions were implemented in the frontend.

(1) **Function of tourist attraction information viewing**

When "View information" is selected on the menu of the top page, "Search from map" and "Search from list" are displayed. Users progress from these to the two pages for this function. Figure 14.2 shows the page for this function when selecting "Search from map". As shown in Fig. 14.2, the categories are temple and shrine, sights and historic spot, nature and scenic spot, park and botanic garden, art museum and museum, and others for tourist attractions, Japanese restaurant, western restaurant, Chinese restaurant and cafe for eateries, and railway station and bus stop for public transportation. Tourist attractions are displayed as pins, and eateries and public transportation are displayed as icons. When users can click or tap on a pin or an icon indicating a tourist attraction, an eatery or public transportation on the digital map, a pop-up containing the details is displayed. Figure 14.3 shows the page for this function when selecting "Search from list". Additionally, this function has two sub-functions developed adopting SNS (the sub-functions of comment submission and tourist attraction bookmark). On this page, using these sub-functions, users can submit comments and bookmark favorite tourist attractions.

(2) **Function of tourist attraction information submission**

When "Submit new tourist attraction" is selected on the menu of the top page, the page for the function of tourist attraction information submission is displayed. Users progress from this to the two pages for the tourist attraction location selection and the tourist attraction information registration in order to submit tourist attraction information. Figure 14.4 shows the page for the tourist attraction location selection. On this page, users can click or tap on any point on the digital map, a pop-up containing the coordinates and the address of the selected location, and a button for switching to the page for the tourist attraction information registration are displayed.



No.	Description
1	Digital map
2	Selected tourist attraction
3	Pop-up containing the details (the name, the number, the category, the address, the business hours, the holiday schedule, the price, the telephone number, the parking space, and the external link (the link to the website) of the selected tourist attraction)
4	"Detail" button
5	“Zoom" button
6	Pins of tourist attraction categories (temple and shrine, sights and historic spot, nature and scenic spot, park and botanic garden, art museum and museum, and others) Icons of eatery categories (Japanese restaurant, western restaurant, Chinese restaurant and cafe) Icons of public transportation categories (railway station and bus)
7	Pop-up containing the list of tourist attraction, the list of eatery and the list of public transportation

Fig. 14.2 Page for the function of tourist attraction information viewing when selecting “Search from map”

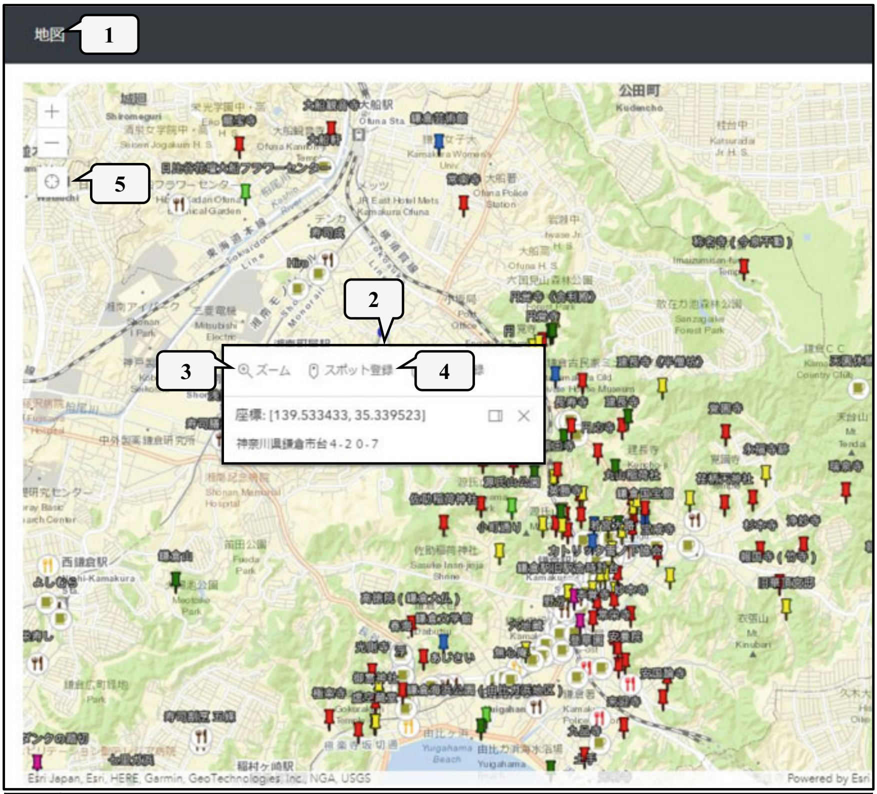
No.	Description
1	Tourist attraction categories (temple and shrine, sights and historic spot, nature and scenic spot, park and botanic garden, art museum and museum, and others)
2	Go to the list of eateries
3	List of tourist attractions
4	List containing the image, the name, the category and the number of registered bookmarks of a tourist attraction
5	"Add bookmark" button

Fig. 14.3 Page for the function of tourist attraction information viewing when selecting "Search from list"

When users click or tap on this button, they progress to this page where they register the details of a tourist attraction.

(3) **Function of tourism congestion display**

When "Tourism congestion map" is selected on the menu of the top page, as shown in Fig. 14.5, the page for the function of tourism congestion display is displayed. On this page, the digital map uses color coding to show the amount of tourism congestion in 44 districts of the operation target area for each month. The "Kamakura tourism congestion" item at the bottom of the digital map can be used to switch the map to



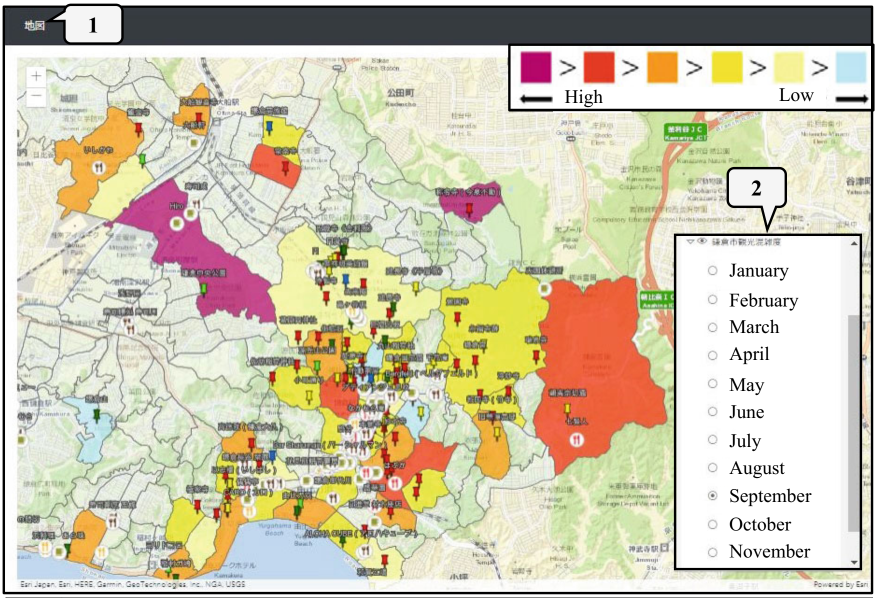
No.	Description
1	Digital map
2	Pop-up containing the coordinates and the address of the selected location
3	“Zoom" button
4	Button for switching to the tourist attraction information registration
5	Button to acquire current location information of the user

Fig. 14.4 Page for tourist attraction location selection

show the tourism congestion in each district for a selected month on the digital map. Users can recognize that congestion district change on a monthly basis using this function. Like “Search from map” of the function of tourist attraction information viewing, when users can click or tap on a pin or an icon indicating a tourist attraction, an eatery or public transportation on the digital map, a pop-up containing the details is displayed.

(4) **Function of tourist attraction recommendation**

When “Tourist attraction recommendations” is selected on the menu of the top page, the page for the function of tourist attraction recommendation is displayed. Table 14.2 shows the relationship between the feature categories of tourist attractions set



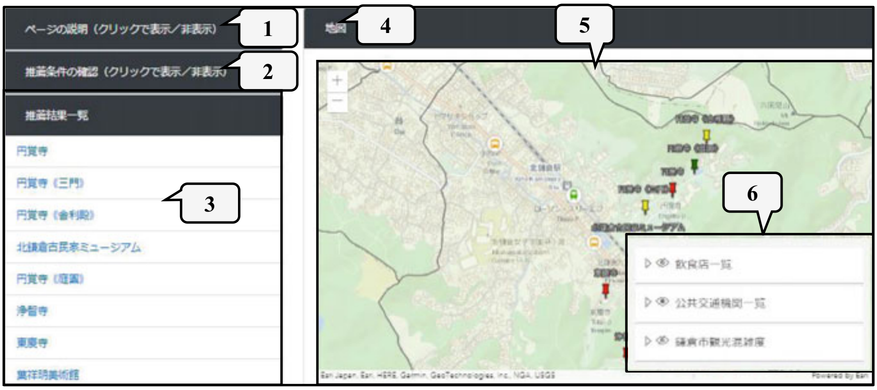
No.	Description
1	Digital map
2	"Kamakura tourism congestion" item containing the list of tourist attraction, the list of eatery, the list of public transportation, and the amount of tourism congestion in each district of the operation target area for each month

Fig. 14.5 Page for the function of tourism congestion display

in Sect. 14.4.3 and users' conditions required for tourist attractions. Regarding the "access (difficult)" and the "access (easy)" conditions, in the operation target area, as the countermeasures against traffic congestion, the use of public transportation is actively promoted, and tourists visit tourist attractions by train and bus. On this page, users indicate the degree of importance placed on each condition required for tourist attractions. This is done by moving a slider on a ten-point scale (1: low importance to 10: high importance). In addition to the above, users select the items of other conditions required for tourist attractions (the "Category", the "Center of recommendation area", the "Distance from center", and the "Maximum number of displayed recommendations"). When users click or tap on the "Make recommendations" button, the page for the recommendation results is displayed. Figure 14.6 shows this page that consists of the condition confirmation, the recommendation results, and the locations of the recommended tourist attractions on the digital map. Like the function of tourism congestion display, the amount of tourism congestion can be shown on the recommendation results on the digital map.

Table 14.2 Relationship between the feature categories of tourist attractions and users' conditions required for tourist attractions

Feature categories of tourist attractions	User's conditions required for tourist attractions
Congestion (high)	Iconic spots
Congestion (low)	Spots that are highly evaluated by tourists but seldom crowded
Access (difficult)	Visit on foot because of inconvenient public transportation (Walk from railway stations or bus stops for more than 20 min)
Access (easy)	Visit via public transportation
Nature tourism	Nature tourism
Historical and cultural tourism	Historical and cultural tourism



No.	Description
1	Explanation of this page
2	Condition confirmation
3	Recommendation results
4	Digital map
5	Locations of the recommended tourist attractions on the digital map
6	Pop-up containing the list of tourist attraction, the list of eatery and the list of public transportation

Fig. 14.6 Page for the recommendation results

14.5.2 System Backend

In the system backend, the following processing for the function of tourist attraction recommendation is conducted.

(1) **Extraction of tourist attractions that match conditions**

The system extracts from the database tourist attractions that match the "Category" specified in the condition and that are within the recommendation area specified in the "Center of recommendation area" and the "Distance from center" conditions.

(2) **Display of the recommendation area on the digital map**

A circle is displayed on the digital map indicating the recommendation area, as set by the items of the "Center of recommendation area" and the "Distance from center" conditions selected by the user.

(3) **Calculation of the degree of similarity**

The degree of similarity between the feature vectors created on the basis of the degree of importance placed by the user on each condition required for tourist attractions and the feature value of a tourist attraction extracted in processing (1) is calculated by applying Eq. (14.1) in Sect. 14.3.2.2. Then, recommendations are issued for the number of tourist attractions specified in the "Maximum number of displayed recommendations" condition.

14.5.3 System Interface

There are two types of system interfaces. One is a PC and mobile information terminal screen interface for users, and the other is a PC screen interface for administrators. The user screen adopts a responsive design, preparing two different types of interfaces depending on the size of the screen of the device that is used to access the system. Administrators can use the administrator page to manage users and the submitted information including images and comments by them. A graphical user interface (GUI) is adopted, which enables administrators to remove malicious users and delete inappropriate information submitted to the system, regardless of the degree of information technology (IT) literacy.

14.6 System Operation

14.6.1 Operation Overview

This system was put in operation for a period of six weeks, for use by people both inside and outside the operation target area. The authors called for people to use the system via the website, Twitter and Facebook of their lab. Users registered when using the system for the first time. Accounts were created by entering their account names, genders, age brackets, and email addresses and passwords. Users entered their email addresses and passwords to log into the system. The system automatically switched to the top page immediately after registration was completed. On my page, users could check and delete submitted information and confirm bookmarked tourist attractions by them. Additionally, they could check and modify their personal information.

14.6.2 Operation Results

Table 14.3 shows the breakdown of users. There were 61 male users and 18 female users, for a total of 79 users. Though the number of male users in their 20 s was particularly high, the user base covered wide generations. During the operation period, information concerning 28 new tourist attraction was submitted. There were 0 comments submitted and 15 tourist attractions were bookmarked by users. However, if the system were operated for a longer period of time, there would be a high potential for more multifaceted use.

14.7 System Evaluation

14.7.1 Overview of the Web Questionnaire Survey for Users

In accordance with the aim of the present study, a web questionnaire survey for users was administered in order to evaluate (1) the usefulness of the key functions and (2) the usability of the system. Thus, the system evaluation was conducted on the basis of users' responses to the web questionnaire survey. Table 14.3 also shows the breakdown of the web questionnaire survey respondents. As shown in Table 14.3, 58 of the 79 users answered the questionnaire survey, and the response rate was 73%.

14.7.2 Evaluation Result

Table 14.4 shows the evaluation results based on the web questionnaire survey for users.

Table 14.3 Breakdown of users and web questionnaire survey respondents

User age (age bracket)	10–19	20–29	30–39	40–49	50–59	60-	Total
Number of users	4	28	18	12	8	9	79
Number of web questionnaire survey respondents	0	21	14	9	7	7	58
Response rate (%)	0.0	75.0	77.8	75.0	87.5	77.8	73.4

Table 14.4 Evaluation results based on the web questionnaire survey for users (%)

Question items	Agree	Somewhat agree	Somewhat disagree	Disagree
Easiness to confirm tourism congestion using the digital map	41.4	43.1	8.6	6.9
Intention to use information concerning congestion periods as a reference and engage in distributed travel	32.8	46.6	17.2	3.4
Intention to use information concerning con-gestation areas as a reference and engage in distributed travel	25.9	56.9	13.8	3.4
Easiness to use the function of tourist attraction recommendation	32.8	48.3	13.8	5.1
Novelty of the recommended tourist attractions	27.6	48.3	15.5	8.6
Interest in recommended tourist attractions	31.0	57.0	10.3	1.7
Effectiveness of the system in providing tourism support	57.0	37.9	3.4	1.7
Wish to use the system in the future	44.9	44.8	8.6	1.7
Hope to engage in distributed travel by utilizing the system	41.4	48.3	8.6	1.7

14.7.2.1 Evaluation of the Usefulness of the Key Functions

(1) Function of tourism congestion display

85% of respondents indicated that they "Agreed" or "Somewhat agreed" that tourism congestion was easy to confirm using the digital map. This shows that displaying tourism congestion on a digital map of Web-GIS was visually intuitive. 79% and 83% of respondents indicated that they "Agreed" or "Somewhat agreed" that they intended to use information concerning congestion periods and areas respectively as a reference and engage in distributed travel. From this, it is clear that the function of tourism congestion display was effective in promoting distributed travel during the sightseeing planning stage. Additionally, four users wished to know the tourism congestion on an hourly basis using this function.

(2) Function of tourist attraction recommendation

81% of respondents indicated that they "Agreed" or "Somewhat agreed" that the function of tourist attraction recommendation was easy to use. Therefore, it would be fair to say that the system design, providing knowledge-based recommendation adopted in this function, and making recommendations when users indicate the degree of importance placed on each condition required for tourist attractions using a ten-point scale from 1 to 10, was effective. 76% of respondents indicated that they "Agreed" or "Somewhat agreed" that they found the recommended tourist attractions novel. Additionally, 88% of respondents indicated that they "Agreed" or "Somewhat agreed" that they were interested in the recommended tourist attractions. Based on the above, it is clear that this function recommended users novel tourist attractions, and stimulated their interests in visiting recommended tourist attractions. However, two users proposed that this function should be improved to propose the appropriate tourist routes to visit the recommended tourist attractions on a digital map of Web-GIS.

14.7.2.2 Evaluation of the Usability of the System

95% of respondents indicated that they "Agreed" or "Somewhat agreed" that the system was effective in providing tourism support. Therefore, it is clear that the system can effectively provide users with tourism support. However, five users suggested that the interfaces of the system should be improved so that they can easily understand the purpose of each function. 90% of respondents indicated that they "Agreed" or "Somewhat agreed" that they wished to use the system in the future. Thus, there is a great deal of promise for ongoing use of the system by the improvement of the functions and the implementation of new ones for better usability for users of wide generations. 90% of respondents indicated that they "Agreed" or "Somewhat agreed" that they hoped to engage in distributed travel by utilizing the system. Based on this, it can be concluded that the system could promote distributed travel by users. Additionally, three users highly evaluated the system and requested that it should be applied to other urban tourist destinations.

14.7.3 Identification of Improvement Measures

The improvement measures of the key functions and the system were identified on the basis of the evaluation results in Sect. 14.7.2 are summarized below.

(1) **Improvement measures for the key functions**

Regarding the function of tourism congestion display, it is necessary to implement the sub-function to automatically infer and display the tourism congestion on an hourly basis in each district of the operation target area on a digital map of Web-GIS, by using reviews submitted to tourism-related web media and adopting artificial intelligence (AI) technology. Additionally, regarding the function of tourist attraction recommendation, it is desirable to add the sub-function to propose the appropriate tourist routes to visit the recommended tourist attractions on a digital map of Web-GIS.

(2) **Improvement measures for the system**

It is indispensable to improve the interfaces by adding the accurate and comprehensible explanations concerning all the functions to the pages of the system. Thus, users can then easily understand the purpose of each function and engage in distributed travel by utilizing the system.

14.8 Conclusion

In the present study, a tourism support system was developed by integrating Web-GIS, recommendation system and SNS with the aim of encouraging distributed travel. The Web-GIS component provided the function of tourism congestion display that indicated the tourism congestion in 44 districts of the operation target area for each month, which was inferred on the basis of 30,356 reviews submitted to tourism-related web media, on a digital map. This provided users with detailed information concerning the timing and locations of congestion in a visually intuitive way, thereby promoting the creation of sightseeing plans that avoid high congestion periods and areas. The recommendation system provided the function of tourist attraction recommendation that adopts knowledge-based recommendation. Feature values of tourist attractions were calculated by utilizing the words extracted from the reviews to recommend users tourist attractions that were novel and tailored to their preferences and tourism purposes, without being swayed by how well-known the tourist attractions were. This encouraged greater dispersion by tourists.

Kamakura City in Kanagawa Prefecture was selected as the operation target area of the system developed in the present study. The system was put in operation for a period of six weeks, for use by people of wide generations both inside and outside this city. During the operation period, there were 79 users and information concerning 28 new tourist attraction was submitted. In order to evaluate the system, a web questionnaire survey for users was administered. 58 users answered the questionnaire survey for the response rate of 73%. Based on the results of the questionnaire survey, it revealed that the function of tourism congestion display can promote tourism that takes congestion periods and areas into consideration. It also showed that the function of tourist attraction recommendation can provide users with novel tourist attraction recommendations and achieve high levels of intent to visit recommended tourist attractions. Thus, it concluded that the system as a whole encouraged distributed travel by users through its utilization.

Regarding future research topics, firstly, when the COVID-19 pandemic is over and people of wide generations will safely go sightseeing, in order to evaluate the usefulness of the key two functions (the functions of tourism congestion display and tourist attraction recommendation) and the usability of the system in detail, it is essential to verify whether users will engage in distributed travel in the operation target area by utilizing the system. Next, based on the system evaluation in Sect. 14.7, it is necessary to make improvements of the system to enhance the usability by adding new sub-functions to the key functions and ameliorating the interfaces. Additionally, it is desirable to advance the societal implementation of the improved system for urban tourist destinations inside and outside Japan.

Acknowledgements In the operation of the tourism support system to promote distributed travel adopting GIS and recommendation system, and the web questionnaire survey of the present study, enormous cooperation was received from those in Japan. We would like to take this opportunity to gratefully acknowledge them.

References

Abe S, Sasaki R, Yamamoto K (2021) Sightseeing support system with augmented reality and no language barriers. In: Geertman S, Pettit C, Goodspeed R, Kauppi A (eds) Lecture notes in urban informatics for future cities. Springer, pp 591–611

An HW, Moon N (2019) Design of recommendation system for tourist spot using sentiment analysis based on CNN-LSTM. J Ambient Intell Humaniz Comput 13:1653–1663

Bali A, Monavari SM, Riazi B, Khorasani M, Zarkeh MMK (2016) A spatial decision support system for ecotourism development in caspian hyrcanian mixed forests ecoregion. Boletim De Ciencias Geodesicas 21(2):340–353

Bongki K (2018) Global tourism promotion and overtourism. J Hiroshima Bunka Gakuen Univ Center Netw Soc 14(1):45–54

Choi K (2020) The occurrence of overtourism and the struggle for sustainable tourism development. KIU J Econ Int Studies 5:193–206

Esmaeili L, Mardani S, Alireza S, Golpayegani H, Madar ZZ (2020) A novel tourism recommender system in the context of social commerce. Expert Syst Appl 149:1–11

García-Palomares JC, Gutiérrez J, Mínguez C (2015) Identification of tourist hot spots based on social networks: a comparative analysis of European metropolises using photo-sharing services and GIS. Appl Geogr 63:408–417

Hida M, Kanaya Y, Kawanaka S, Matsuda Y, Nakamura Y, Suwa H, Fujimoto M, Arakawa Y, Yasumoto K (2020) On-site trip planning support system based on dynamic information on tourism spots. Smart Cities 3(2):212–231

Himono D, Ieiri Y, Hishiyama R (2020) Analysis of guided person's behavior history at the time of guidance using AR-POI. Proceeding of the 82th national convention of IPSJ 2020(1):227–228

Hirota M, Endo M, Kato D, Ishikawa H (2019) Discovering hotspots using photographic orientation and angle of view from social media site. Int J Inf Soc 10(3):109–117

Ikeda T, Yamamoto K (2014) Development of social recommendation GIS for tourist spots. Int J Adv Comput Sci Appl 5(12):8–21

Itou J, Mori T, Munemori J, Babguchi N (2018) Development of a stroll support system using Route display on a map and image sharing service. In: Egi H, Yuizono T, Baloian N, Yoshino T, Ichimura S, Rodrigues A (eds) Collaboration technologies and social computing. Springer, pp 48–55

Japan Tourism Agency, Ministry of Land, Infrastructure, Transport and Tourism (2019) Toward a sustainable advanced country of tourism. 46 p

Jannach D, Zanker M, Felferning A, Friedrich G (2010) Recommender systems: An introduction. Cambridge University Press, 352 p

Kamakura City (2016) The third stage Kamakura City tourism basic plan. 78 p

Kamishima T (2008) Algorithms for recommender systems (2). J Japanese Soc Artif Intell 23(1):89–103

Kato Y, Yamamoto K (2020) Sightseeing spot recommendation system that takes into account the visiting frequency of users. Geo-Information 9(7):411.https://doi.org/10.3390/ijgi9070411

Kaufmann M, Siegfried P, Huck L, Stettler J (2019) Analysis of tourism hotspot behaviour based on geolocated travel blog data: the case of qyer. Int J Geo-Inf 8(11):493. https://doi.org/10.3390/ijgi8110493

Kawai Y, Kumamoto T (2018) A system to visualize regional congestion degree based on the analysis result of swarm tweets. DEIM Forum 2018:3–4

Koga Y, Yamamoto K (2021) A sightseeing planning support system with gamification. J Geogr Inf Syst 13(4):485–507

Kurashima T, Iwata T, Irie G, Fujimura K (2010) Travel route recommendation using geotags in photo sharing sites. Proceedings of the conference on information and knowledge management (CIKM), 579–588

Kurata Y (2012) Potential-of-interest maps for mobile tourist information services. In: Fuchs M, Ricci F, Cantoni L (eds) Information and communication technologies in tourism 2012. Springer, pp 239–248

Kurata Y, Ai H, Sanada F (2015) Creation of innovative tourist maps based on the user-posted data of a photo-sharing site. Int J Tourism Sci 8:151–154

Ma Q (2019) Utilization and analysis of user generated contents toward personalized and distributed sightseeing. Trans Inst Syst Control Inf Eng 63(1):32–37

Majid A, Chen L, Chen G, Mirza HT, Hussain I, Woodward J (2013) A context-aware personalized travel recommendation system based on geotagged social media data mining. Int J Geogr Inf Sci 27(4):662–684

Majid A, Chen L, Tura HT, Hussain I, Chen G (2015) A system for mining interesting tourist locations and travel sequences from public geo-tagged photos. Data Knowl Eng 95:66–86

Masron T, Mohamaed B, Marzuki A (2015) GIS base tourism decision support system for Langkawi Island, Kedah, Malaysia. Theoret Empirical Res Urban Manag 10(2):21–35

Masron T, Ismail N, Marzuki A (2016) The conceptual design and application of web-based tourism decision support systems. Theoret Empirical Res Urban Manag 11(2):64–75

Mizutani Y, Yamamoto K (2021) A sightseeing spot recommendation system that takes into account the change in circumstances of users. Int J Geo-Inf 6(10):303. https://doi.org/10.3390/ijgi6100303

Noguera JM, Barranco MJ, Segura RJ, Martinez L (2012) A mobile 3D-GIS hybrid recommender system for tourism. Inf Sci 215:37–52

Singh SP, Sharma J, Singh P (2011) A geo-referenced information system for tourism (GeoRIST). Int J Geomatics Geosci 2(2):456–464

Sugimoto K (2018) Use of GIS-based analysis to explore the characteristics of preferred viewing spots indicated by the visual interest of visitors. Landsc Res 43(3):345–359

Sumitomo Y, Ishino T, Kubo K, Yue W (2020) Construction and demonstration experiment of similar spot recommendation system based on emotion word contained in word-of-mouth information. J Japan Personal Comput Appl Technol Soc 14(1):29–35

Totsuka K, Murata T (2021) Evaluation of an avoiding method for tourist concentration using agent-based simulation. Proceeding of the 24th subcommittee workshop of social system of the SICE section of system and information, 116–120

United Nations World Tourism Organization (UNWTO) (2016) Overtourism?—Understanding and managing urban tourism growth beyond perceptions, 12 p

Zhuang C, Ma Q, Liang X, Yoshikawa M (2015) Discovering obscure sightseeing spots by analysis of geo-tagged social images. Proceedings of the 2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), 590–595

Chapter 15
Applying the AURIN Walkability Index at the Metropolitan and Local Levels by Sex and Age in Australia

Arsham Bassiri Abyaneh, Andrew Allan, Johannes Pieters, Sekhar Somenahalli, and Ali Soltani

Abstract Apart from the ever-growing body of research aiming to develop novel walkability assessment tools, conducting systematic literature reviews on these planning support tools has been emerging as a new research theme in the field of urban analytics. While the walkability index of the Australian Urban Research Infrastructure Network (AURIN) has been used in several studies in the last decade, there is still a need for comprehensive application and correlation analyses of this tool at different geographical scales and with various socio-demographics and walking behaviours as covariates. To analyse the level of responsiveness of this tool, this study applies it at two different national and metropolitan scales comparing the walkability of 175 Local Government Areas in all States and Territories as well as 3047 neighbourhoods in the Greater Adelaide Metropolitan Area. Although the body of knowledge criticises using large units of analysis in the walkability assessments, the AURIN Walkability Index (AWI) is more representative at the national level. Furthermore, in each unit of analysis, the walk-to-work behaviour of females and the youngest age groups demonstrate the strongest associations with the walkability scores.

A. Bassiri Abyaneh (✉) · A. Allan · J. Pieters
UniSA Creative, University of South Australia, City West Campus, Adelaide, SA 5000, Australia
e-mail: arsham.bassiri_abyaneh@mymail.unisa.edu.au

A. Allan
e-mail: Andrew.Allan@unisa.edu.au

J. Pieters
e-mail: Johannes.Pieters@unisa.edu.au

S. Somenahalli
UniSA STEM, University of South Australia, Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia
e-mail: Sekhar.Somenahalli@unisa.edu.au

A. Soltani
UniSA Business, University of South Australia, City West Campus, Adelaide, SA 5000, Australia
e-mail: Ali.Soltani@unisa.edu.au

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_15

Keywords Australian Urban Research Infrastructure Network (AURIN) · Walkability Index · Planning support system · Geographic Information System (GIS) · Australia

15.1 Introduction

Walking as one of the healthiest, cheapest, easiest, most accessible, sustainable, democratic and equitable modes of transportation presents a vast range of social, ecological, political and economic advantages (Abastante et al. 2020). In the last decade, walkability has been one of the most trending topics in the fields of public health and urban planning (Shashank and Schuurman 2019). The fast-growing sedentary lifestyle and significant dependence on motorised vehicles have resulted in poor quality of life outcomes with associated health risks. According to Guthold et al. (2018), globally, 25% or more of adults fail to reach the recommended threshold of 150-min of moderate-intensity weekly physical activity to maintain good health. This is particularly problematic in Australia, where sedentary activity is a by-product of Australia's reputation as one of the world's most car-dependent nations. For example, more than 65% of Australians aged +15-year use private cars (either as drivers or passengers) solely for their daily travel to work or in combination with other transport modes (Australian Bureau of Statistics 2016).

A vast range of quantitative assessment tools, generally called walkability indices, have been developed in recent years to analyse the level of walkability at different scales from pathways and streets to neighbourhoods and districts. Advancements in the field of Geographic Information Systems (GIS) have resulted in an ever-growing availability and application of urban datasets in walkability studies (Wang and Yang 2019).

However, a knowledge gap exists in terms of the appropriate scale and unit of analysis for GIS-based walkability indices as well as comparing how different groups of citizens with various personal and socio-demographic characteristics value walkability (Fancello et al. 2020). Most studies have just focussed on a narrow and specific group of people such as the elderly, school children or females without conducting any comparative analyses between different categories of pedestrians (Gaglione et al. 2022; Lee et al. 2020; Golan et al. 2019). In addition, they employ a wide range of methodologies, in terms of walkability indicators and geographical units of analyses rendering comparison difficult.

Australian Urban Research Infrastructure Network (AURIN) which is funded by the Australian Government is an online platform providing planners, researchers and urban analysts with the technical infrastructure to enable evidence-based policymaking in Australian cities with access to a network of planning support tools and datasets (Giles-Corti et al. 2014). Although the AURIN Walkability Index (AWI) has been applied in several walkability studies during the recent decade (Giles-Corti et al. 2014; Bassiri Abyaneh et al. 2021), there is still a need for comprehensive analyses of this planning support tool in terms of the walking behaviours of various categories of pedestrians and its effectiveness at different geographical scales.

This chapter attempts to address the following three research questions:

(a) Is there a correlation between the AWI scores (independent variable) and the census "walking to work" data (dependent variable) across the Australian States and Territories?
(b) What are the gender and age gaps in walking to work within the Australian context?
(c) Is there a consistent correlation pattern between the walkability scores at the national level and the metropolitan level?

Therefore, in order to understand the latest trends in the fast-moving field of GIS-based walkability tools and indices, a systematic review of literature emerging over the past five years was completed covering the most recent walkability criteria, socio-demographic profiling, scales and units of analysis as well as methods of visualisation. Subsequently, the AWI was applied at the national scale to compare the walkability of Australian Metropolitan Areas in all States and Territories. As a result of this applied approach, both local and metropolitan jurisdictions in Australia were objectively ranked in terms of their relative walkability levels. Afterwards, this tool was applied at the metropolitan scale by analysing Greater Adelaide as the main case study. Finally, age- and gender-based correlation analyses of walk-to-work behaviour were conducted to allow comparisons between AURIN's walkability scores and the behaviour of different genders and ages.

15.2 Literature Review

In recent decades, there has been an emerging research theme globally in systematically categorising the development and application of GIS-based walkability assessment tools (and their input factors), to solve urban planning scenarios at different spatial scales (Wang and Yang 2019; Blečić et al. 2020; Shields et al. 2021).

To identify the relevant and recent publications regarding the GIS-based walkability assessment indices and their assessed walkability criteria, a systematic literature review was conducted utilising the PRISMA method, which is a systematic data collection protocol involving four steps, namely identification, screening, eligibility and inclusion (Fig. 15.1). In the first step, the Scopus database was chosen for conducting keyword searches because of its inclusivity in comparison with other search engines such as the Web of Science and Google Scholar (Aghaei Chadegani et al. 2013). Three categories of keywords were applied to operate the TITLE-ABS-KEY procedure in the Scopus search engine including (1) ["walkability"/"walk"], (2) ["GIS"/"Geographic Information System"], and (3) ["index"/"tool"/"formula"/"measur*"/"assess"/"evaluation"] which led to the retrieval of 663 records. Secondly, to identify the most advanced and innovative GIS-based walkability assessment tools, the range of the publication year was set to

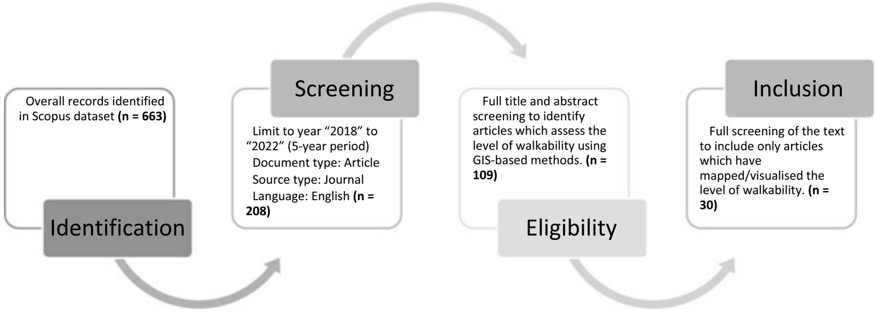


Fig. 15.1 Steps of the PRISMA protocol for the article collection process (*Note* n represents the number of articles)

the last five years (2018–2022). Furthermore, to ensure consistency and homogeneity, only English peer-reviewed journal articles were considered. In the third step, a full screening of the titles and abstracts was conducted to exclude less focused publications. In the final step, as the main inclusion criteria, only publications were chosen for critical review which visualised and mapped walkability in their text.

15.2.1 Included Criteria and Attributes

There are a vast range of walkability criteria and indicators factored in the reviewed articles. Based on the operational definitions of walkability, two main groups of indicators are witnessed in the literature including environmental (statistical)-based indicators and accessibility-based indicators (Horak et al. 2022).

The first group mostly calculates the local urban characteristics and conditions using ratios without considering any specific distances. The majority of these studies used area-based measures including intersection density, population density, public transit stop density, entropy score of the land-use mix, retail and service density, lighting density, the Kernel density of recreation and retail facilities, etc. (Bödeker et al. 2018; Ribeiro and Hoffimann 2018; Al Shammas and Escobar 2019; Molina-García et al. 2019; Kenyon and Pearce 2019; Sabzali Yameqani and Alesheikh 2019; Shashank and Schuurman 2019; Lee et al. 2020; Bartzokas-Tsiompras and Photis et al. 2021; Telega et al. 2021; Adams et al. 2022; Gaglione et al. 2022; Lam et al. 2022; Shashank et al. 2022). Although all the reviewed publications were published in the last five years, most of their walkability attributes were based on the classic concept of so-called 3Ds (Density, Diversity and Design) developed by Cervero and Kockelman (1997). This concept was further extended to 5Ds by adding "Destination accessibility" and "Distance to transit" (Ewing and Cervero 2010).

The second group is a category of measures that depends highly on the distribution of and the distance from the destinations and evaluates travel parameters in a road network to selected destinations which emphasise walkability as a proxy for amenity availability (Sun et al. 2021; Zhao et al. 2020; Horak et al. 2022). The most frequently used method of calculating walkability in these studies is the distance decay function from several categories of land uses or amenities (Kim et al. 2019; Sun et al. 2021; Amaya et al. 2022) which is similar to the Walk Score algorithm (Zhang and Mu 2019). The logic behind Walk Score as an accessibility-based evaluation system has been widely used in the walkability literature.

Apart from the above-mentioned groups of indicators, advancements in the field of spatial sciences have led to the incorporation of walkability tools of elements such as tree canopy coverage data, the Urban Heat Island (UHI) indicator, the level of Greenness (Normalised Difference Vegetation Index (NDVI)), the Land Surface Temperature data and the effects of building shades on sidewalks, as integral to determining walking comfort levels in urban settings (Taleai and Yameqani 2018; Al Shammas and Escobar 2019; Motieyan et al. 2022; Rahman 2022).

15.2.2 Socio-Demographic Profiling

According to Blečić et al. (2020), an important requirement for the evaluation of the walkability tools and studies is the researchers' intention for considering and comparing the socio-demographic categories of users at various stages of the research design. This can result in an improved understanding of the preferences, attitudes and needs of different groups of people with diverse characteristics and abilities when choosing a walking route in an urban area.

However, half of the reviewed articles did not consider any specific profiling of the users and solely focused on the built environment features. It is worth mentioning that, during the article collection process, discussions of the different socio-demographic profiling were noted in several walkability publications in the field of public health where there was no visualisation or mapping.

A number of reviewed studies refer to walkability from the perspective of specific segments of society, such as the elderly (Gaglione et al. 2022; Horak et al. 2022), women (Golan et al. 2019) and school children (Lee et al. 2020). By contrast, other articles consider a mixture of socio-demographic characteristics such as age, gender, education, occupation, sports routine, ethnic background, household income and neighbourhood socio-economic status, mostly as covariates (Bödeker et al. 2018; Fancello et al. 2020; Zhang and Mu 2019; Lam et al. 2022).

15.2.3 Scale of the Analysis/Geography

The scale and the geography of the walkability analyses vary from university campuses (Zhang and Mu 2019; Abastante et al. 2020) and neighbourhoods (Gaglione et al. 2022; Amaya et al. 2022) to cities (Taleai and Yameqani 2018; Jabbari et al. 2018; Shashank and Schuurman 2019; Telega et al. 2021) and even nationally (Lam et al. 2022). However, most of the studies were conducted at the metropolitan scale (Bartzokas-Tsiompras and Photis 2020; Deng et al. 2020; Delclòs-Alió et al. 2019; Ribeiro and Hoffimann 2018). Surprisingly, there is only one study from Australia in the reviewed articles which analysed the City of Sydney Local Government Area (about 27 km^2) as a sole case study (Rahman 2022).

15.2.4 Unit of Analysis

According to Horak et al. (2022), choosing the unit of analysis is the first step in measuring walkability in any study. It impacts the detail of the visualisation and the readability of the results, particularly by non-experts (Blečić et al. 2020). The spatial units of analysis in the reviewed articles vary from administrative neighbourhoods and census tracts (Bödeker et al. 2018; Molina-García et al. 2019; Shashank and Schuurman 2019; Bartzokas-Tsiompras and Photis 2020; Lam et al. 2022; Motieyan et al. 2022) to 400–2000-m network buffer from individual addresses (Kenyon and Pearce 2019; Deng et al. 2020; Lee et al. 2020; Adams et al. 2022). Furthermore, at the micro-level scales, the spatial units range from a square grid of 5–500-m pixels (Delclòs-Alió et al. 2019; Kim et al. 2019; D'Orso and Migliore 2020; Zhang and Mu 2019; Amaya et al. 2022; Horak et al. 2022; Shashank et al. 2022) to network/street segments (Jabbari et al. 2018; Gaglione et al. 2022; Rahman 2022) and even nodes (Fancello et al. 2020).

However, Horak et al. (2022, p. 4) assert that the choice of administrative units which are mostly representatives of the policy regulations of different jurisdictions as the units of analysis in the walkability studies could be problematic. They mentioned several typical issues including "abrupt temporal changes, variability in size, Modifiable Area Unit Problem (MAUP) and strong border effects".

15.2.5 Method of Visualisation

Several methods of visualisation of the results have been witnessed in the reviewed articles including Polygons (Shashank and Schuurman 2019), Rasters (Telega et al. 2021) and Lines (Gaglione et al. 2022). According to Blečić et al. (2020), the creative visualisation of walkability data will significantly expand the useability of the walkability assessment methods beyond professional urban analysts and researchers to users with basic spatial science skills such as politicians.

Some of the reviewed walkability studies offer a visualisation method that differs from their units of analysis. For instance, while Taleai and Yameqani (2018) analysed their assessed walkability indicators at the 20 and 30-m buffers around every road section, they, eventually, allocated the scores to the road sections and visualised them as lines.

15.3 Methodology

The AURIN portal provides access to Australia's most homogenous and comprehensive urban datasets making comparative analyses possible between different case studies at different scales almost anywhere in Australia. In this chapter, since our walkability analyses were at a strategic/regional level and we did not need to calculate the distance decay function for various locations, we solely considered the environmental (statistical)-based groups of indicators which can be calculated objectively by the AURIN Walkability Index (AWI), particularly its "Land-use Mix Entropy Measure". In terms of the socio-demographic profiles of pedestrians, based on the only publicly available walking dataset of Australia which can be accessed from the AURIN portal, we narrowed our focus to different age and gender groups as covariates. Another rationale for choosing the AWI is its flexibility in terms of the scales and units of analysis. Eventually, the walkability scores of the AWI can easily be visualised as polygons for each unit of analysis.

15.3.1 Study Area

In order to evaluate the capabilities of the AWI and its level of responsiveness, this study was conducted at two different geographical scales, at the national and metropolitan levels. Australia's census statistical geography boundaries and categorisation provide a basis for the application of the AWI planning support tool. As can be seen in Fig. 15.2, the smallest census tracts are called Mesh Blocks (MB) which contain 30 to 60 dwellings. Australian Bureau of Statistics (ABS) (2022) broadly determines land uses for all Mesh Blocks including residential, commercial, industrial, parkland, etc. From the agglomeration of whole Mesh Blocks, Statistical Areas Level 1 (SA1s) were produced with populations from 200 to 800 persons. The SA1s are designed to represent the detailed population and housing data at the neighbourhood level while maintaining confidentiality. Statistical Areas Level 2 (SA2s) are medium-sized areas built up from the whole SA1s with a population between 3,000 and 25,000 persons. They represent suburbs within cities and catchments of rural areas. Thereafter, from aggregating the SA2s, Statistical Areas Level 3 (SA3s) are provided which match the administrative boundaries of "cities" and approximate Local Government Areas (LGAs). Subsequently, the SA3s' assemblage generates Statistical Area Level 4 (SA4) with populations above 100,000 persons. Eventually, SA4s produce States, Territories and the whole of Australia.

In this study, we focused on SA3s and SA1s as the units of analysis for conducting the walkability assessment using the AURIN walkability index at the national and metropolitan scales, respectively. These statistical areas were selected since the census data of method of travel to work as the only nationwide dataset of walking behaviour of Australians is publicly available at these geographical boundaries.

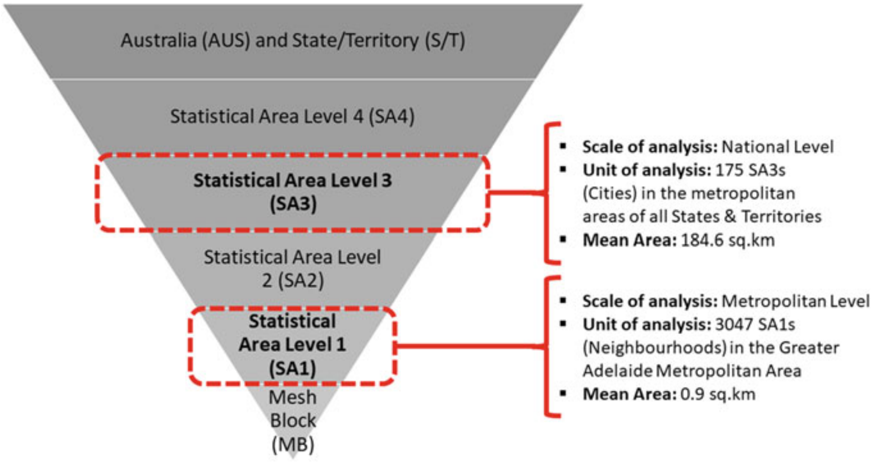


Fig. 15.2 The characteristics of the case studies within the hierarchy of Australia's census geography boundaries

At the national level, 175 metropolitan SA3s in all 8 States and Territories of Australia were chosen. For maintaining consistency, the SA3s with areas of larger than 1400 km^2 were eliminated. The mean area of these chosen SA3s is 184.6 km^2. The smallest SA3 with a 10.6 km^2 area is Leichhardt in New South Wales (NSW) and the largest SA3 is 1388.1 km^2 for Adelaide Hills in South Australia (SA). As the boundaries of the SA3s are closely aligned with the Local Government Areas, the findings can be used as a walkability performance indicator for Local Governments and as a basis for further research comparing the level of implementation of active travel policies in different jurisdictions and from a broader strategic planning viewpoint.

At the metropolitan level, 3047 SA1s in the Greater Adelaide metropolitan area were selected. The inclusion criteria for these SA1s were, firstly, location within the Greater Adelaide Metropolitan area and, secondly, a residential population. The mean area of the SA1s is 0.9 km^2.

15.3.2 AURIN Walkability Index (AWI)

The AURIN portal offers several urban and geographic analysis tools among which the "Walkability Index Within Areas" tool has been used in this study. AURIN (2021) calculates the three compound elements of the urban fabric including connectivity, land use mix and gross population density to analyse the walkability of each unit of analysis as the walking catchments. These indicators are the classic criteria of walkability based on the "3Ds" concept which has been used in the majority of the reviewed articles. The AURIN walkability formula, as well as a brief description of the components' calculating method, are shown below:

$$\text{AURIN's Walkability Index} = [\text{Z-score Connectivity}] + [\text{Z-score Land-use mix}] + [\text{Z-score Population density}]$$

where:

- *Z-score Connectivity = The z-score of the total number of three (or more) way street intersections (connections) per square kilometre in the walking catchments;*
- *Z-score Land-use mix = The z-score of the "Land-use Mix Entropy Measure" which measures the extent to which there is an equal distribution of each land use within the catchments;*
- *Z-score Population density = The z-score of the average population per hectare for each of the catchments.*
- *Z-score* $X = \frac{x_i - x}{s}$ *where* X_i *is the non-normalised score of observation i;* $\overline{X}$ *is the sample mean and s is the sample standard deviation.*

In terms of the inputs for the connectivity measure, the Street Network dataset released by PSMA Australia Limited (2020) was used. Also, the dataset of 2016 Usual Residential Population and Dwelling Count at the Mesh Block level (ABS 2017) was used to calculate the population density as well as the land-use mix measure considering five categories of 'Commercial', 'Education', 'Hospital/Medical', 'Parkland' and 'Residential'.

15.3.3 Correlation of the AWI with Observed Data of "Walking to Work"

As one of the questions of the 2016 Census of Population and Housing by ABS, respondents were asked to indicate their method of travel to work. They were able to select one or more methods. Up to three methods were recorded and the results for the + 15-year-old population (active population) were released at the scales of SA1s, SA2s, SA3s and SA4s. According to the ABS (2021), if a person walked some of the way to work and used other methods, it is not included as an additional method. For example, if they walked and then caught the bus, then only the "Bus" mode would be selected. Hence, at both scales of analyses, for the active population of each SA3 and SA1, all methods of commute to work which have at least one walking component including "Train", "Bus", "Ferry", "Tram", "Train and Bus", "Train and Ferry", "Train and Tram", "Bus and Ferry", "Bus and Tram" as well as "Walked only" have been summed to calculate a measure of "walking to work" to represent each unit of analysis in the correlation process.

To analyse and compare the level of association of different components of the AURIN walkability index and the "walking to work" behaviour of different age and gender groups, the 2016 census datasets of 'SA3-G59 Method of Travel to

Work by Sex' and 'SA3-W22 Method of Travel to Work by Sex by Age' for the national scale and the dataset of 'SA1-G59 Method of Travel to Work by Sex' for the metropolitan scale were used. The Spearman correlation test was conducted comparing the percentage of the population aged 15 years or older of each unit of analysis walking to work for each gender and age group with every measured walkability indicator using IBM SPSS Statistics ver. 28. To maintain reliability, two SA3s and thirty-two SA1s with the highest overall walkability scores were excluded as the outliers in this step based on the inter-quartile range (IQR) rule.

15.4 Analysis and Findings

15.4.1 At the National Level

Figure 15.3 illustrates the level of the walkability of 175 SA3s in the eight Greater metropolitan areas of Australia. The majority of the Central Business Districts (CBDs) are relatively walkable in comparison with the outer suburbs. According to Table 15.1, in each SA3 and at the national level, the Spearman correlation coefficients indicate the AWI is a good predictor of "walking to work" behaviour. In addition, as can be seen in Figs. 15.4 and 15.5, the scatterplots of the relationship between the percentage of active persons (15-years or older) in each SA3 who walk to work whether as a single mode or as a combination with public transport modes show that the dissimilarities between groups of different ages and genders are relatively striking. Walking to work is more strongly associated with walkability scores for females than for males. Furthermore, the walking behaviour of two age groups of 15–24 and 25–34 years highlights the most significant associations with the objective walkability z-scores. However, the association is significantly lower in older groups.

15.4.2 At the Metropolitan Level

Since ABS has not released the 2016 census data on age-based travel to work at the SA1 level, the correlation analysis at the metropolitan level was based only on different gender groups. Figure 15.6 illustrates that the Adelaide CBD, Mawson Lakes and Glenelg can be considered the most walkable neighbourhoods of this metropolitan area. By contrast, most of the outer suburbs are heavily car-dependent. Interestingly, in Adelaide's original metropolitan 2010 Planning Strategy, these areas were designated Transit Oriented Developments (latterly "Activity Centres") (South Australian Government, Department of Planning and Local Government 2010); hence, there was a conscious planning effort to develop these areas as walkable precincts. In addition, according to Fig. 15.7, although the AURIN walkability scores are positively correlated with the walking-to-work behaviour of both gender groups,

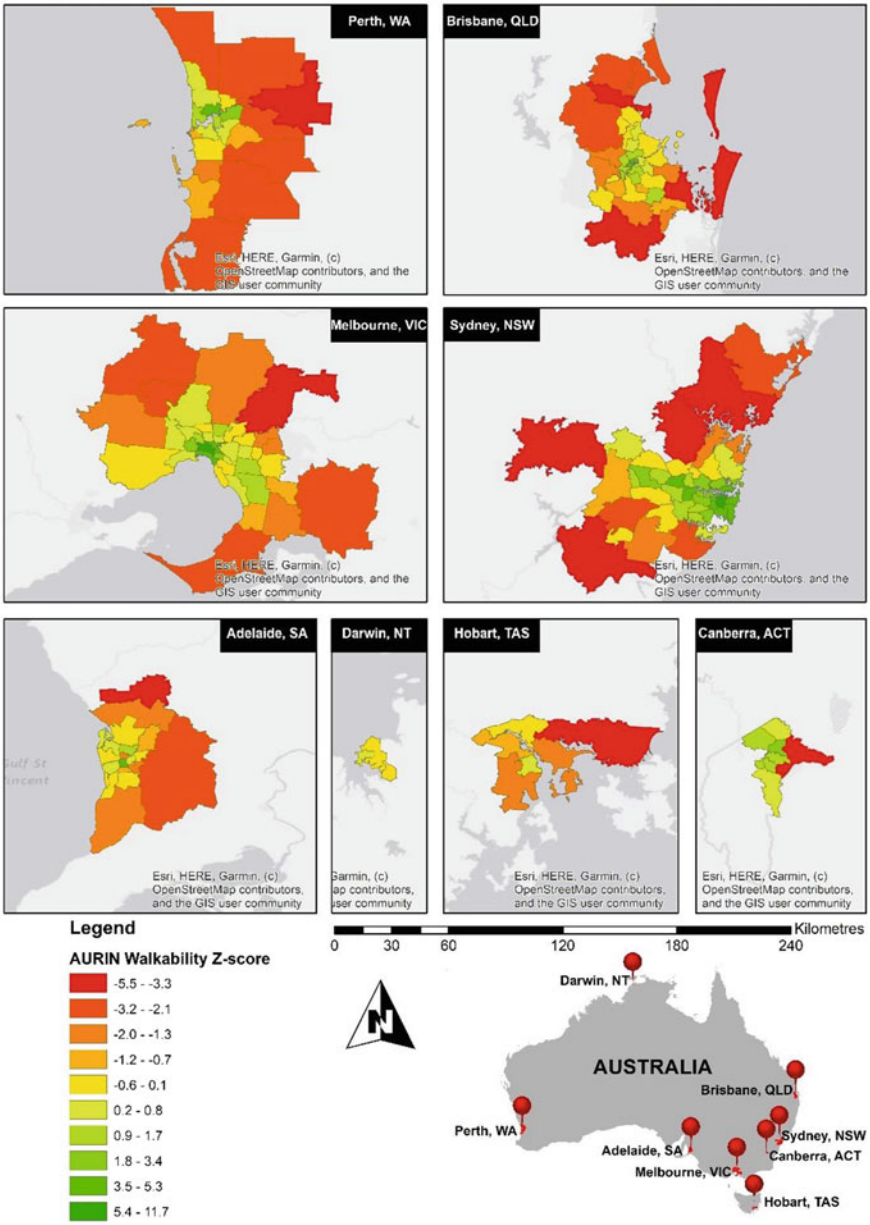


Fig. 15.3 The national-level walkability analysis in all Greater metropolitan areas of Australia

Table 15.1 The correlation coefficients of the Spearman test for different gender and age groups

Spearman's rho	Gender		Age							Persons
	Female	Male	15–24	25–34	35–44	45–54	55–64	65–74	75+	
Connectivity z-score	0.751	0.719	0.689	0.711	0.657	0.592	0.527	0.358	0.223	0.740
Land-use mix z-score	0.220	0.183	0.173	0.197	0.130	0.094	0.059	−0.050	0.016	0.200
Population density z-score	0.890	0.838	0.783	0.792	0.739	0.682	0.635	0.542	0.370	0.871
Overall walkability z-score	0.748	0.703	0.665	0.698	0.638	0.574	0.527	0.369	0.255	0.729

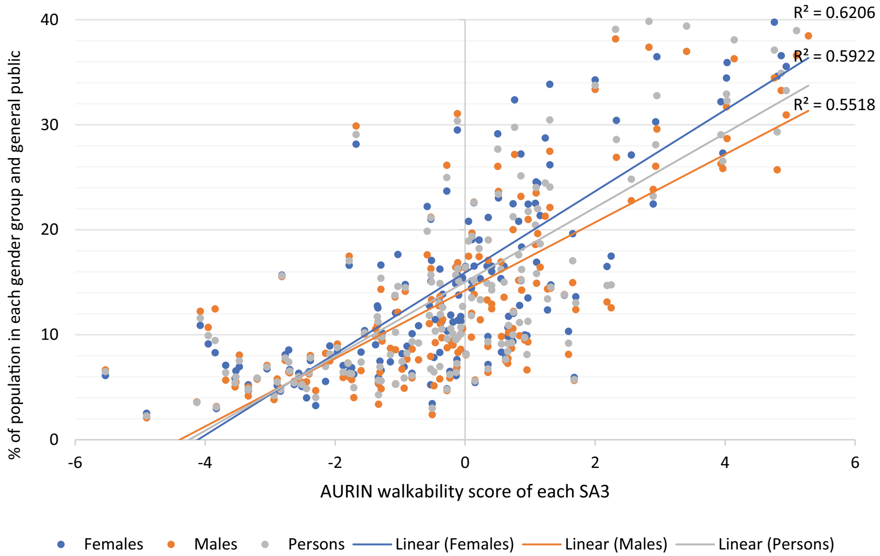


Fig. 15.4 The scatterplot of the relationship between the percentage of the active population walking to work for both gender groups in each SA3 and the walkability score at the national scale

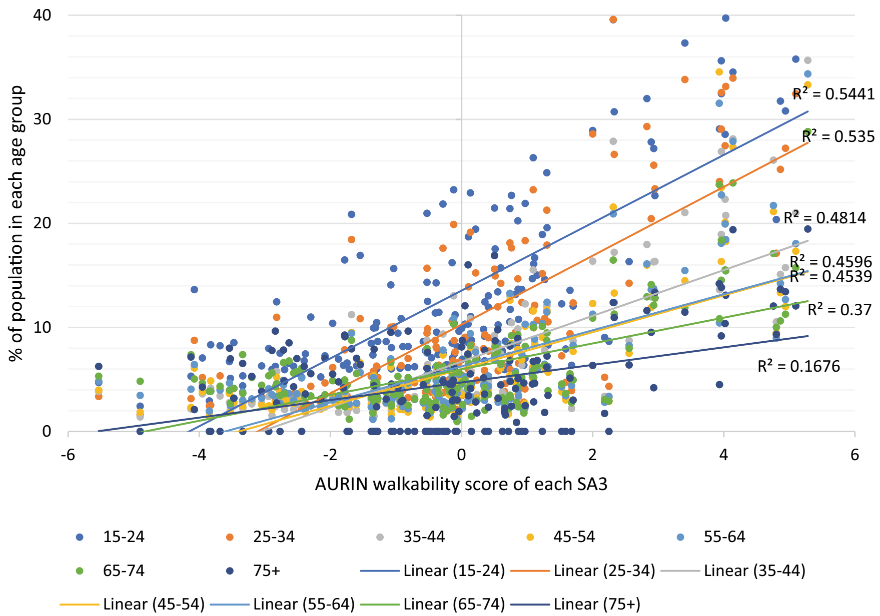


Fig. 15.5 The scatterplot of the relationship between the percentage of the active population walking to work in seven different age groups in each SA3 and the walkability score at the national level

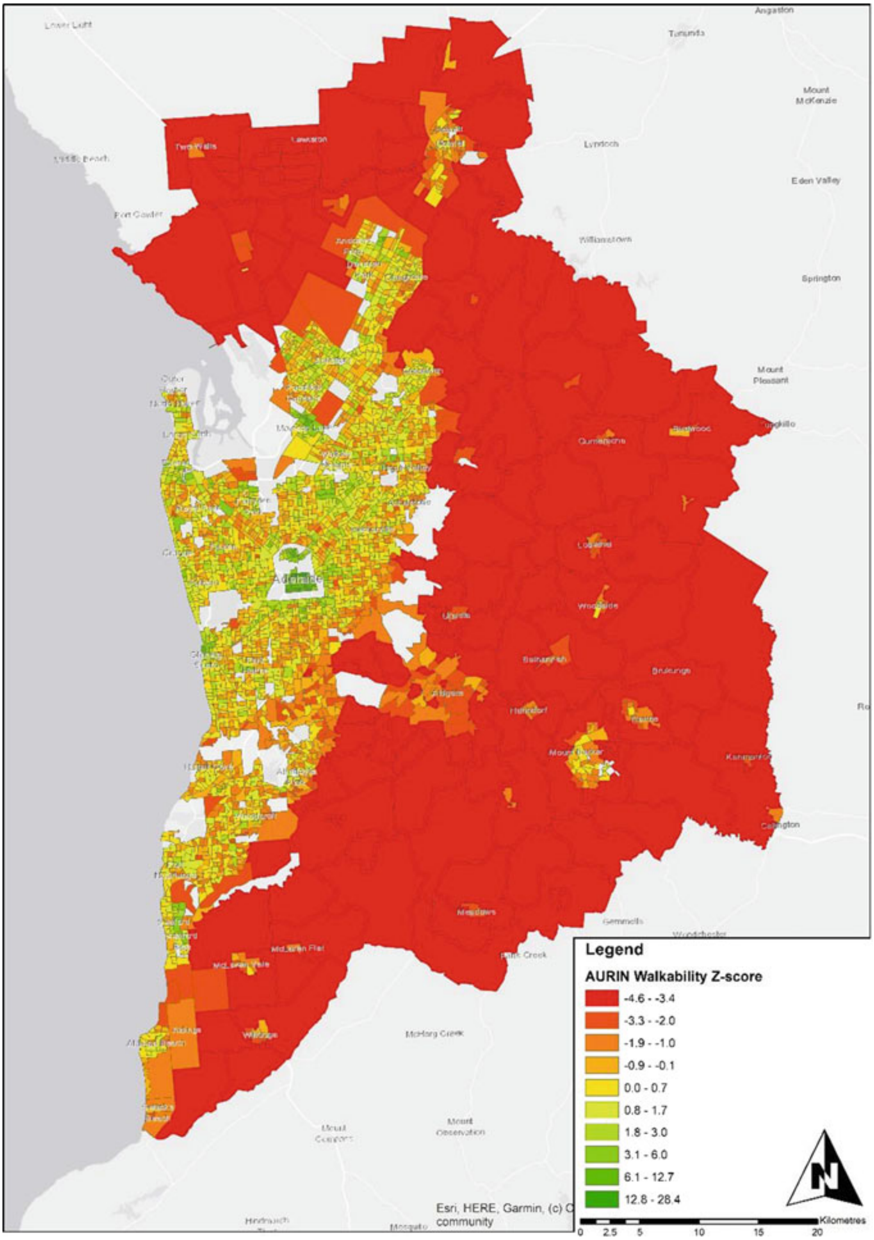


Fig. 15.6 The metropolitan-level walkability analysis in the Greater Adelaide

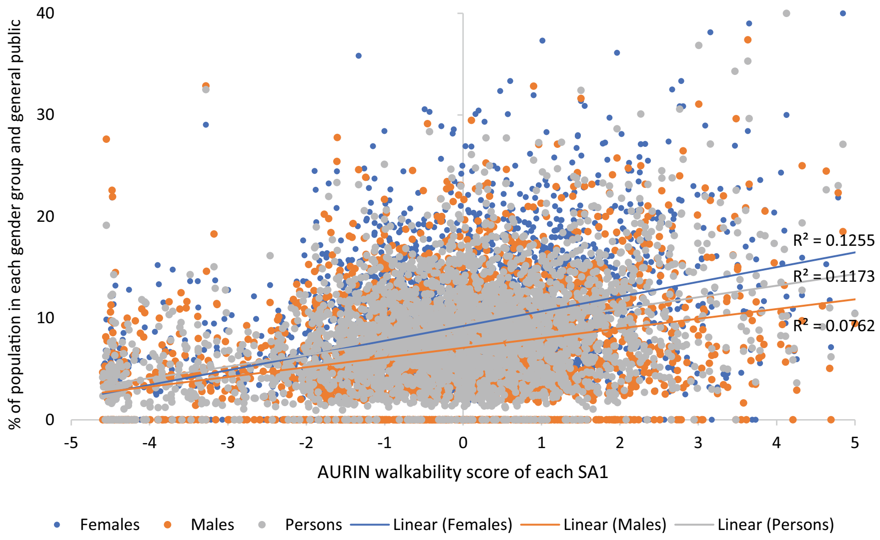


Fig. 15.7 The scatterplot of the relationship between the percentage of the active population walking to work for both gender groups in each SA1 and the walkability score at the scale of the Greater Adelaide Metropolitan Area

the size of this association is relatively small in comparison with the national-level analyses. This indicates that with the current accuracy of the national datasets of land-use and street networks on the AURIN portal, the AWI is more suitable for conducting analyses on larger units such as SA3s and from a more holistic and strategic viewpoint.

15.5 Discussion and Conclusion

It is important to recognise that people have different perspectives toward walking which might be a necessity for those who do not have access to other transportation options (Shields et al. 2021). Currently, the research trend on walkability is a shift from homogeneous definitions of pedestrians to taking ethnicity, gender and sexual diversity into account, as well as ergonomic differences by age and ability (Wang and Yang 2019; Shields et al. 2021).

One of the main planning instruments available via the AURIN portal is its walkability toolkit which has been used in this study to assess the level of walkability in two different scales and for a variety of age groups and genders. Regarding the first research question, a key finding is that the AWI results which comprised from calculating the most-frequently cited walkability indicators of ‘Connectivity’, ‘Land-use mix’ and ‘Population density’ are significantly correlated with the only publicly available national walk-to-work data in each unit of analysis.

With regards to the second research question and at both scales of the analyses, a higher correlation coefficient exists between walkability scores and walking to work behaviour among females in comparison with males. This gap is evident in the findings from several previous walkability studies (Yang et al. 2022). For instance, Adeel et al. (2017) indicate that, generally, females tend to engage in walking more than males. Similarly, a later study in India, revealed that women achieve higher levels of physical activity than men (Adlakha and Parra 2020). In addition, a Slovakian walking study revealed that females place a greater emphasis on safety aspects than males (Rišová and Sládeková Madajová 2020). According to two gender-specific walkability analyses in Iran, females are more sensitive to walking distances on both work and school trips (Hatamzadeh et al. 2020; Javadpoor et al. 2023).

Another important outcome of this study is the remarkable dissimilarity between the level of associations of walkability scores and the 'walk-to-work' behaviour of each age group. Generally, younger age groups are more likely to walk to work if their surroundings are conducive to this type of behaviour. According to Akinci et al. (2022), generally, older adults, for instance, are less involved in walking than younger adults due to physical limitations. In addition, apart from the 3Ds, the walk-to-work behaviour of older age groups might be more dependent on the micro-level built environmental features, such as poor lighting, stairs, slopes and damaged pavements which might make walking difficult (Garvin et al. 2012).

In terms of addressing the third and final research question, to our knowledge, this is the first study to conduct a comparative strategic walkability analysis between different jurisdictions in all eight States and Territories of Australia. The outcomes of the national-level analyses can be used by different Local Governments and strategic planners to understand the walkability performance of their respective jurisdictions. The aim of the analysis was to compare all 175 SA3s of varying sizes, based on their location, including Central Business Districts, inner, middle and outer suburbs, as well as regional areas, using the 'Quantiles' method, ranking them from the highest to the lowest walkability scores (Table 15.2). On the other hand, and for SA1s at the Adelaide metropolitan scale, the association of walkability scores and the walk-to-work behaviours for each age group is smaller compared to national-level analyses which might be the result of the Modifiable Areal Unit Problem (MAUP) (Gerell 2016).

These results show that socio-demographic factors can significantly affect walking behaviour and should be considered in future walkability analyses. The collection of more detailed socio-demographic factors beyond gender when capturing walking data should be considered an important direction for future studies on the current national walking datasets and the AURIN walkability toolkit. For instance, the collection and visualisation of the walking data based on ethnicity, level of fitness, disability status and trip purposes can be suggested.

One of the main limitations of this study is its narrow focus on the 'walk-to-work' data and not considering other walking purposes such as recreational walking and walking for errands. We cannot therefore claim that this measure has provided the best assessment of the AWI. This highlights that a remarkable gap exists in the available comprehensive walking datasets in Australia. The ABS data only covers

Table 15.2 The rankings of the walkability scores of all analysed SA3s categorised into four size groups

<31.5 sq.km	31.5-66.5 sq.km	66.5-172.4 sq.km	>172.4 sq.km
Sydney Inner City	Auburn	Mount Druitt	Tullamarine - Broadmeadows
Melbourne City	Parramatta	Bankstown	Richmond - Windsor
Brisbane Inner	Perth City	Belconnen	Brighton
Yarra	North Canberra	Monash	Wyndham
Port Phillip	Strathfield - Burwood - Ashfield	Dandenong	Penrith
Leichhardt	Ryde - Hunters Hill	Kingston	Rockingham
Eastern Suburbs - North	Chatswood - Lane Cove	Canning	Playford
Adelaide City	Belmont - Victoria Park	Port Adelaide - West	Campbelltown (NSW)
Eastern Suburbs - South	St Marys	Gungahlin	Hobart - North East
Marrickville – Sydenham – P	South Canberra	Joondalup	Onkaparinga
Botany	Hurstville	Stirling	Melton - Bacchus Marsh
North Sydney - Mosman	Cottesloe - Claremont	Warringah	Hobart - South and West
Stonnington - West	Blacktown	Tuggeranong	Casey - South
Kogarah - Rockdale	Darebin - North	Brimbank	Whittlesea - Wallan
Canada Bay	Port Adelaide - East	Cockburn	Mandurah
Maribyrnong	Springwood - Kingston	Blacktown - North	Mornington Peninsula
Manly	Merrylands - Guildford	Fairfield	Wanneroo
Woden Valley	Hobart Inner	Rocklea - Acacia Ridge	Swan
Weston Creek	Melville	Baulkham Hills	Serpentine - Jarrahdale
South Perth	Keilor	Darwin Suburbs	Cardinia
Brisbane Inner - North	Cronulla - Miranda - Caringbah	Knox	Caboolture
Canterbury	Darwin City	Mitcham	Adelaide Hills
Holland Park - Yeronga	Chermside	Camden	The Hills District
Carlingford	West Torrens	Ku-ring-gai	Bringelly - Green Valley
Sherwood - Indooroopilly	Hobsons Bay	Mt Gravatt	Sunbury
Brisbane Inner - West	Boroondara	Salisbury	Sutherland - Menai - H
Nathan	Whitehorse - West	Wynnum - Manly	Wyong
Essendon	Sandgate	Springfield - Redbank	Kalamunda
Brunswick - Coburg	Marion	Liverpool	Bribie - Beachmere
Stonnington - East	Charles Sturt	Nundah	Armadale
Darebin - South	Bayside	Tea Tree Gully	Macedon Ranges
Carindale	Strathpine	Casey - North	Mundaring
Brisbane Inner - East	North Lakes	Gosnells	Gawler - Two Wells
Norwood - Payneham – St P	Bayswater - Bassendean	Loganlea - Carbrook	Nillumbik - Kinglake
Moreland - North	Bald Hills - Everton Park	Hobart - North West	Cleveland - Stradbroke
Whitehorse - East	Manningham - West	Frankston	Redcliffe
Pennant Hills - Epping	Banyule	Browns Plains	Sorell - Dodges Ferry
Campbelltown (SA)	Glen Eira	Kwinana	Canberra East
Centenary	Palmerston	Pittwater	Gosford
Holdfast Bay	Forest Lake - Oxley	Hornsby	Blue Mountains
Unley	Fremantle	Capalaba	Wollondilly
Sunnybank	Maroondah	The Gap - Enoggera	Jimboomba
Burnside	Beenleigh	Kenmore - Brookfield - Moggill	Dural - Wisemans Ferry
Prospect - Walkerville	Manningham - East	Narangba - Burpengary	

Australian Capital Territory · New South Wales · Northern Territory · Queensland · South Australia · Tasmania · Victoria · Western Australia

the ‘walk-to-work’ behaviour and is published for the limited socio-demographic profiles including gender, age and occupation groups. Another limitation is the lack of accurate data on the other aspects of the built environment such as the effects of terrain, perceptions of crime, perceptions of safety from moving vehicles, perceptions of environmental comfort (e.g. availability of shelters, shades, etc.), perceptions of physical comfort (e.g. availability of public amenities, etc.) and perceived visual attractiveness which might significantly be associated with the walking behaviour (Bassiri Abyaneh et al. 2021).

In conclusion, with the everyday emergence of interactive methods of accessing and manipulating the urban data on urban dashboards and platforms such as the AURIN portal, the need for further review and study of the GIS-based walkability indices regarding their scales and units of analysis as well as methods of visualisation will be more evident.

Acknowledgements This research is supported by a UniSA International Graduate Research Scholarship.

References

Abastante F, Lami IM, Riccia LL, Gaballo M (2020) Supporting resilient urban planning through walkability assessment. Sustainability (Basel, Switzerland) 12(19):8131. https://doi.org/10.3390/su12198131

Adams MA, Phillips CB, Patel A, Middel A (2022) Training computers to see the built environment related to physical activity: detection of microscale walkability features using computer vision. Int J Environ Res Public Health 19:4548. https://doi.org/10.3390/ijerph19084548

Adeel M, Yeh AGO, Zhang F (2017) Gender inequality in mobility and mode choice in Pakistan. Transportation (dordrecht) 44(6):1519–1534. https://doi.org/10.1007/s11116-016-9712-8

Adlakha D, Parra D (2020) Mind the gap: gender differences in walkability, transportation and physical activity in urban India. J Transp Health 18:100875. https://doi.org/10.1016/j.jth.2020.100875

Aghaei Chadegani A, Salehi H, Md Yunus M et al (2013) A comparison between two main academic literature collections: web of science and Scopus databases. Asian Soc Sci 9(5):18–26. https://doi.org/10.5539/ass.v9n5p18

Akinci ZS, Delclos-Alio X, Vich G et al (2022) How different are objective operationalizations of walkability for older adults compared to the general population? A systematic review. BMC Geriatrics 22(1):1–673. https://doi.org/10.1186/s12877-022-03233-x

Al Shammas T, Escobar F (2019) Comfort and time-based Walkability index design: a GIS-based proposal. Int J Environ Res Public Health 16(16):2850. https://doi.org/10.3390/ijerph16162850

Amaya V, Moulaert T, Gwiazdzinski L, Vuillerme N (2022) Assessing and qualifying neighborhood walkability for older adults: construction and initial testing of a multivariate spatial accessibility model. Int J Environ Res Public Health 19(3):1808. https://doi.org/10.3390/ijerph19031808

AURIN (2021) Walkability Index Within Areas—AURIN. Australian Urban Research Infrastructure Network. https://aurin.org.au/resources/workbench-user-guides/portal-user-guides/analysing-your-data/walkability-tools/walkability-index-within-areas/. Accessed 15 Jun 2021

Australian Bureau of Statistics (2016) 2016 Census All persons QuickStats. https://www.abs.gov.au/census/find-census-data/quickstats/2016/0/. Accessed 27 Sep 2022

Australian Bureau of Statistics (2017) ABS—Usual Residential Population and Dwelling Count (MB) 2016. https://portal.aurin.org.au/. Accessed 25 Jun 2021

Australian Bureau of Statistics (2021) Method of travel to work (MTWP). https://www.abs.gov.au/census/guide-census-data/census-dictionary/2021/variables-topic/transport/method-travel-work-mtwp. Accessed 20 Sep 2022

Australian Bureau of Statistics (2022) Census geography glossary. https://www.abs.gov.au/census/guide-census-data/geography/census-geography-glossary. Accessed 2 Oct 2022

Bartzokas-Tsiompras A, Photis Y (2020) Does neighborhood walkability affect ethnic diversity in Berlin? insights from a spatial modelling approach. Europ J Geogr 11(1):163–187. https://doi.org/10.48088/ejg.a.bar.11.1.163.187

Bassiri Abyaneh A, Allan A, Pieters J, Davison G (2021) Developing a GIS-based tourist walkability index based on the aurin walkability toolkit—case study: sydney CBD. In: Geertman SCM, Pettit C, Goodspeed R, Staffans A (eds) Urban informatics and future cities. The Urban Book Series. Springer, Cham, pp 233–256. https://doi.org/10.1007/978-3-030-76059-5_13.

Blečić I, Congiu T, Fancello G, Trunfio GA (2020) Planning and design support tools for walkability: a guide for urban analysts. Sustainability 12:4405. https://doi.org/10.3390/su12114405

Bödeker M, Finne E, Kerr J, Bucksch J (2018) Active travel despite motorcar access. A city-wide, GIS-based multilevel study on neighborhood walkability and active travel in Germany. J Transp Health 9:8–18. https://doi.org/10.1016/j.jth.2018.03.009

Cervero R, Kockelman K (1997) Travel demand and the 3DS: density, diversity, and design. Transp Res Part D: Transp Environ 2(3):199–219. https://doi.org/10.1016/s1361-9209(97)00009-6

Delclòs-Alió X, Gutiérrez A, Miralles-Guasch C (2019) The urban vitality conditions of Jane Jacobs in Barcelona: Residential and smartphone-based tracking measurements of the built environment in a Mediterranean Metropolis. Cities 86:220–228. https://doi.org/10.1016/j.cities.2018.09.021

Deng C, Dong X, Wang H et al (2020) A data-driven framework for walkability measurement with open data: A case study of triple cities, New York. ISPRS Int J Geo Inf 9(1):36. https://doi.org/10.3390/ijgi9010036

D'Orso G, Migliore M (2020) A GIS-based method for evaluating the walkability of a pedestrian environment and prioritised investments. J Transp Geogr 82:102555. https://doi.org/10.1016/j.jtrangeo.2019.102555

Ewing R, Cervero R (2010) Travel and the built environment. J Am Plann Assoc 76(3):265–294. https://doi.org/10.1080/01944361003766766

Fancello G, Congiu T, Tsoukiàs A (2020) Mapping walkability. A subjective value theory approach. Socio-Economic Planning Sciences 72:100923. https://doi.org/10.1016/j.seps.2020.100923

Gaglione F, Gargiulo C, Zucaro F (2022) Where can the elderly walk? A spatial multi-criteria method to increase urban pedestrian accessibility. Cities 127:103724. https://doi.org/10.1016/j.cities.2022.103724

Garvin T, Nykiforuk CIJ, Johnson S (2012) Can we get old here? seniors' perceptions of seasonal constraints of neighbourhood built environments in a northern, winter city. Geografiska Annaler, Series B, Human Geography 94(4):369–389. https://doi.org/10.1111/geob.12004

Gerell M (2016) Smallest is better? the spatial distribution of arson and the modifiable areal unit problem. J Quant Criminol 33(2):293–318. https://doi.org/10.1007/s10940-016-9297-6

Giles-Corti B, Macaulay G, Middleton N et al (2014) Developing a research and practice tool to measure walkability: A demonstration project. Health Promot J Austr 25(3):160–166. https://doi.org/10.1071/he14050

Golan Y, Henderson J, Lee N, Weverka A (2019) Gendered walkability: building a daytime walkability index for women. Journal of Transport and Land Use 12(1):501–526. https://doi.org/10.5198/jtlu.2019.1472

Guthold R, Stevens GA, Riley LM, Bull FC (2018) Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants. Lancet Global Health 6(10):e1077–e1086. https://doi.org/10.1016/s2214-109x(18)30357-7

Hatamzadeh Y, Habibian M, Khodaii A (2020) Walking mode choice across genders for purposes of work and shopping: a case study of an Iranian city. Int J Sustain Transp 14(5):389–402. https://doi.org/10.1080/15568318.2019.1570404

Horak J, Kukuliac P, Maresova P et al (2022) Spatial pattern of the walkability index, walk score and walk score modification for elderly. ISPRS Int J Geo Inf 11(5):279. https://doi.org/10.3390/ijgi11050279

Jabbari M, Fonseca F, Ramos R (2018) Combining multi-criteria and Space syntax analysis to assess a pedestrian network: the case of Oporto. J Urban Des 23(1):23–41. https://doi.org/10.1080/13574809.2017.1343087

Javadpoor M, Soltani A, Fatehnia L, Soltani N (2023) How the built environment moderates gender gap in active commuting to schools. Int J Environ Res Public Health 20(2):1131. https://doi.org/10.3390/ijerph20021131

Kenyon A, Pearce J (2019) The socio-spatial distribution of walkable environments in Urban Scotland: a case study from glasgow and edinburgh. SSM—Population Health 9:100461. https://doi.org/10.1016/j.ssmph.2019.100461

Kim EJ, Won J, Kim J (2019) Is Seoul walkable? assessing a walkability score and examining its relationship with pedestrian satisfaction in Seoul. Korea. Sustain 11(24):6915. https://doi.org/10.3390/su11246915

Lam T, Wang Z, Vaartjes I et al (2022) Development of an objectively measured walkability index for the Netherlands. Int J Beha Nutrition Phys Activity 19(1):1–50. https://doi.org/10.1186/s12966-022-01270-8

Lee S, Lee C, Nam JW et al (2020) School walkability index: application of environmental audit tool and GIS. J Transp Health 18:100880. https://doi.org/10.1016/j.jth.2020.100880

Molina-García J, Menescardi C, Estevan I et al (2019) Neighborhood built environment and socioeconomic status are associated with active commuting and sedentary behavior, but not with leisure-time physical activity, in university students. Int J Environ Res Public Health 16(17):3176. https://doi.org/10.3390/ijerph16173176

Motieyan H, Kaviari F, Mostofi N (2022) Quantifying walking capability: a novel aggregated index based on spatial perspective and analyses. Pap Reg Sci 101(2):483–503. https://doi.org/10.1111/pirs.12652

PSMA Australia Limited (trading as Geoscape Australia) (2020) PSMA—Transport & Topography—Street Network (Line) August 2020. https://portal.aurin.org.au/. Accessed 25 Jun 2021.

Rahman A (2022) A GIS-based, microscale walkability assessment integrating the local topography. J Transp Geogr 103:103405. https://doi.org/10.1016/j.jtrangeo.2022.103405

Ribeiro AI, Hoffimann E (2018) Development of a neighbourhood walkability index for Porto Metropolitan Area: how strongly is walkability associated with walking for transport? Int J Environ Res Public Health 15(12):2767. https://doi.org/10.3390/ijerph15122767

Rišová K, Sládeková Madajová M (2020) Gender differences in a walking environment safety perception: a case study in a small town of Banská Bystrica (Slovakia). J Transp Geogr 85:102723–102728. https://doi.org/10.1016/j.jtrangeo.2020.102723

Sabzali Yameqani A, Alesheikh AA (2019) Predicting subjective measures of walkability index from objective measures using artificial neural networks. Sustain Cities Soc 48:101560. https://doi.org/10.1016/j.scs.2019.101560

Shashank A, Schuurman N (2019) Unpacking walkability indices and their inherent assumptions. Health Place 55:145–154. https://doi.org/10.1016/j.healthplace.2018.12.005

Shashank A, Schuurman N, Copley R, Lear S (2022) Creation of a rough runnability index using an affordance-based framework. Environ Planning B: Urban Analyt City Sci 49(1):321–334. https://doi.org/10.1177/23998083211003886

Shields R, Gomes da Silva EJ, Lima e Lima T, Osorio N (2021) Walkability: a review of trends. Journal of Urbanism 1–23.https://doi.org/10.1080/17549175.2021.1936601

South Australian Government, Department of Planning and Local Government (2010) The 30-Year Plan for Greater Adelaide.

Sun Y, Lu W, Sun P (2021) Optimization of walk score based on street greening—a case study of Zhongshan Road in Qingdao. Int J Environ Res Public Health 18(3):1277. https://doi.org/10.3390/ijerph18031277

Taleai M, Yameqani AS (2018) Integration of GIS, Remote Sensing and multi-criteria evaluation tools in the search for healthy walking paths. KSCE J Civ Eng 22(1):279–291. https://doi.org/10.1007/s12205-017-2538-x

Telega A, Telega I, Bieda A (2021) Measuring walkability with GIS—methods overview and new approach proposal. Sustainability 13(4):1883. https://doi.org/10.3390/su13041883

Wang H, Yang Y (2019) Neighbourhood walkability: a review and bibliometric analysis. Cities 93:43–61. https://doi.org/10.1016/j.cities.2019.04.015

Yang Y, Sasaki K, Cheng L, Liu X (2022) Gender differences in active travel among older adults: non-linear built environment insights. Transportation research. Part D Transp Environ 110:103405. https://doi.org/10.1016/j.trd.2022.103405.

Zhang X, Mu L (2019) The perceived importance and objective measurement of walkability in the built environment rating. Environ Planning B: Urban Analyt City Sci 47(9):1655–1671. https://doi.org/10.1177/2399808319832305

Zhao J, Sun G, Webster C (2020) Walkability scoring: why and how does a three-dimensional pedestrian network matter? Environ Planning B: Urban Analyt City Sci 48(8):2418–2435. https://doi.org/10.1177/2399808320977871

Chapter 16
Predicting Urban Heat Island Mitigation with Random Forest Regression in Belgian Cities

Mitali Yeshwant Joshi, Daniel G. Aliaga, and Jacques Teller

Abstract An abundance of impervious surfaces like building roofs in densely populated cities make green roofs a suitable solution for urban heat island (UHI) mitigation. Therefore, we employ random forest (RF) regression to predict the impact of green roofs on the surface UHI (SUHI) in Liege, Belgium. While there have been several studies identifying the impact of green roofs on UHI, fewer studies utilize a remote-sensing-based approach to measure impact on Land Surface Temperatures (LST) that are used to estimate SUHI. Moreover, the RF algorithm, can provide useful insights. In this study, we use LST obtained from Landsat-8 imagery and relate it to 2D and 3D morphological parameters that influence LST and UHI effects. Additionally, we utilise parameters that influence wind (e.g., frontal area index). We simulate the green roofs by assigning suitable values of normalised difference-vegetation index and built-up index to the buildings with flat roofs. Results suggest that green roofs decrease the average LST.

Keywords Green roofs · Random forest regression · Urban heat island (UHI) · Land surface temperature (LST)

16.1 Introduction

Unprecedented urban growth has led to increased building densities resulting in limited green spaces in cities, exacerbating the impacts of climate change (Dong et al. 2020; Wang et al. 2022). Consequently, urban areas are experiencing higher

M. Y. Joshi (✉) · J. Teller
LEMA Research Group, Urban & Environmental Engineering Department, University of Liège, Liege, Belgium
e-mail: mjoshi@uliege.be

J. Teller
e-mail: Jacques.Teller@uliege.be

D. G. Aliaga
Department of Computer Science, Purdue University, West Lafayette, IN, USA
e-mail: aliaga@cs.purdue.edu

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_16

temperatures as compared to rural regions, which is known as the urban heat island (UHI) effect (Stewart and Oke 2012; Santamouris 2013). Negative impacts of UHI effect include increased energy consumption, carbon emissions, decreased comfort levels, and global warming (Bowler et al. 2010). Thus, to circumvent these impacts, identifying solutions to mitigate the UHI effect is crucial.

The UHI phenomenon is primarily caused by a high density of built-up areas, as well as low albedo of urban surfaces, resulting in absorption of excess solar radiation (Razzaghmanesh et al. 2016). In this scenario, green roofs are highly relevant owing to the abundance of building rooftops made of impervious surfaces in cities (Francis and Jensen 2017; Joshi and Teller 2021). Green roofs prevent the absorption of short-wave radiation and act as thermal insulators (Razzaghmanesh et al. 2016). Green roofs also prevent heat from entering the structures in summer, reducing energy consumption (di Giuseppe and D'Orazio 2014). Moreover, green roofs also increase evapotranspiration and natural ventilation within built-up areas (Li et al. 2014). Along with this, green roofs also have other benefits in terms of biodiversity and runoff retention. Thus, considering their multifold benefits, green roofs are a suitable strategy for UHI mitigation (Jamei et al. 2021).

Researchers have analyzed green roofs and their impact using numerous methods. Most of the studies use numerical modelling, simulations and statistical analysis to analyze greening scenarios (Bartesaghi Koc et al. 2018). However, micro-scale urban canopy models (UCM) and microclimate simulation models like ENVImet cannot be employed for entire city due to computational demands (Mirzaei, 2015; Lin et al. 2021). While mesoscale models like weather research and forecasting (WRF) coupled with UCM aid in conducting research at a regional scale, running these models at high resolution require significant computational resources. Therefore, most of the WRF studies are carried out at a resolution of 1 km (Yang and Bou-Zeid 2019; Wang et al. 2022). Moreover, analyzing and processing these models are generally challenging for most urban planners (Lin et al. 2021). Therefore, remote sensing approach can be advantageous as finer resolution datasets are increasingly available. Furthermore, the analysis using remote sensing can be straightforward with existing geographic information system (GIS) softwares like QGIS and ArcGIS for urban practitioners (Bartesaghi Koc et al. 2018; Lin et al. 2021).In this study, we employ a remote sensing-based approach for analyzing the impact of green roofs on UHI.

Traditionally, UHI effect is classified into surface and air UHI, referring to the surface and air temperature respectively (Roth et al. 1989; Kleerekoper et al. 2012; Kim and Brown 2021). The impact of Green roofs on air temperature is debatable, however, their impact on surface temperature is significant (Berardi et al. 2014; Francis and Jensen 2017). Thus, we focus mainly on the impact of green roofs on surface temperature. This estimate, referred to as Land Surface Temperature (LST) is obtained using satellite images.

UHI effect is the outcome of dynamic interactions between the macroclimate and urban morphology (Boccalatte et al. 2020). Thus, along with greening, several other morphological parameters influence LST. The relationship of LST with several morphological parameters has been investigated using satellite data (El-Zeiny and

Effat 2017). For example, the Normalised difference vegetation index (NDVI) has a strong to moderate negative correlation with LST, whereas normalized difference built-up index (NDBI) has a positive correlation (Adeyeri et al. 2017; Govil et al. 2020). Parameters such as building densities also influence LST. 3D data was also strongly influencing LST in some studies (Zha et al. 2010). Moreover, detailed 3D parameters such as frontal area index (FAI) and sky view factor (SVF), can be now computed in a raster format for facilitating a better analysis (Asadi et al. 2020). Subsequently, 3D parameters can represent the contribution of shadows, solar radiation and orientation.

The relationship between LST and related parameters have been previously analyzed using ordinary least square (OLS) regression and graphically weighted regression (Deilami et al. 2018). Recent studies also have used machine learning in predicting changes in LST (Jato-Espino et al. 2022; Lyu et al. 2022). For example, Asadi et al. (2020) use Artificial neural network (ANN) to predict changes in LST after implementing green roofs in Austin, Texas. Their study showed that LST decreased by 1.96 degree Celsius on an average after greening 3.2% of the roofs. The study used 2D and 3D urban morphological parameters in their analysis. Although the parameters are comprehensive for the task, parameters influencing wind flow, such as frontal area index (FAI) could have enhanced the analysis (Wang et al. 2021). Moreover, random forest (RF) regression has proven to be robust in predicting several scenarios (Jato-Espino et al. 2022; Lyu et al. 2022). Along with this, RF regression prediction is regarded as being unaffected by the multicollinearity and distribution of data (Matsuki et al. 2016; Busato et al. 2023). Thus, in this study, we explore the performance of RF regression in predicting changes in LST caused by green roofs.

16.2 Methodology

Figure 16.1 represents the broad methodology used in this study.

16.3 Study Area and Dataset

In this study, we analyze the Brussels capital region and the city of Liege in Belgium. Brussels is the capital of Belgium and has an area of 161.4 km^2 with around 1,222,637 inhabitants (Christis et al. 2019). As it is the national capital, it is highly dense and compact with limited space for developing green infrastructure within the city. Liege, on the other hand, is located in the Wallonia region of Belgium. The city is the third largest city in terms of population in the country, with an area of 69 km^2 and 196,296 inhabitants. The city is highly compact with significant building density in the center and open residential areas in the outskirts (Joshi et al. 2022). Thus, roof greening is a suitable UHI mitigation strategy for both cities.

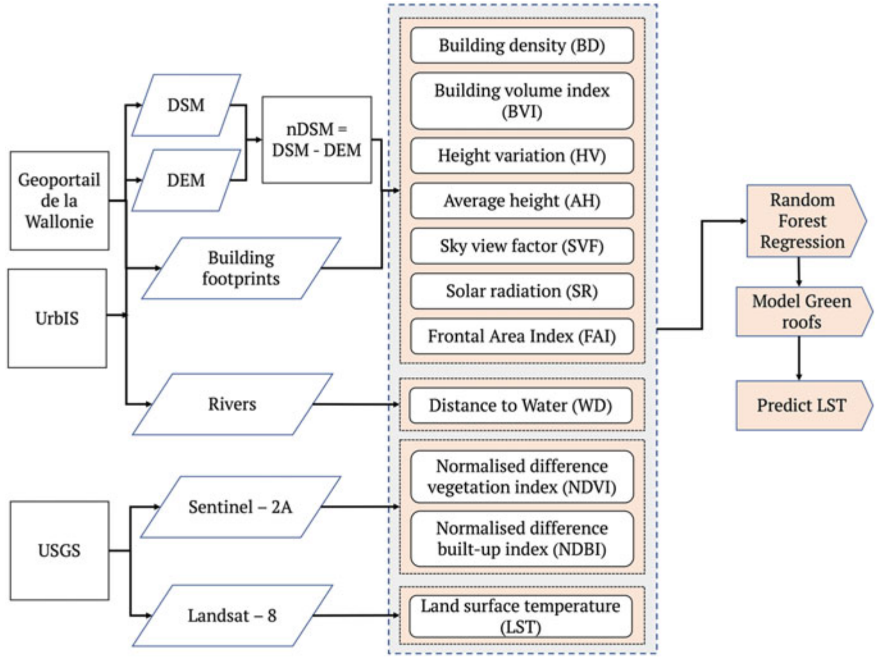


Fig. 16.1 Methodology

We use four datasets mainly for computing the parameters influencing LST. Table 16.1 presents the dataset used and respective sources for Liege and Brussels.

We processed all rasters to 30 m resolution as it is the resolution of LST obtained from Landsat-8.

Table 16.1 Datasets used for analysis with their respective sources for Liege and Brussels

Datasets	Liege	Brussels
Building footprints	PICC (Projet Informatique de Cartographie Continue) dataset, with an accuracy of less than 25 cm	The footprints are obtained from UrbIS online (data platform for Brussels capital region)
Building heights	Digital surface model (DSM) and the digital elevation model (DEM)	3D model of Brussels from UrbIS
LST	Landsat-8 level 1 image captured on July 18, 2021	
NDVI and NDBI	Sentinel-2 multispectral image obtained on 21st July 2021	

16.4 Parameters Influencing LST

16.4.1 Building Density

To obtain building density (BD), we first transform the building footprints to a raster with 1 m resolution. Thereafter, we aggregate the raster by summing the building pixels to a raster with 30 m resolution, representing building density. Thus, building density is computed as:

$$BD = \sum_{i=1}^{n} a_i \tag{16.1}$$

where a_i is the 1 by 1 m pixel covered by building and n is the total number of 1 m pixels in 30 m pixel.

16.4.2 Building Volume Index

Building volume index (BVI) is the building volume in a pixel and it is calculated as follows:

$$BVI = \sum_{i=1}^{n} a_i \times h_i \tag{16.2}$$

where a_i is the 1 m pixel covered by building, h_i is the height of the building in the 1 m pixel and n is the total number of 1 m pixels in 30 m pixel.

16.4.3 Sky View Factor

Sky view factor (SVF) is the ratio of proportion of sky visible from the ground at a given position, to the proportion of sky not obstructed by the surrounding built-up (Rodler and Leduc 2019). We calculate it with the Relief Visualisation Toolbox of QGIS 3 (Zakšek et al. 2011; Kokalj and Somrak 2019).

We use the DSM and the building footprint dataset to generate the raster with building height information. We consider open spaces and roads along with the bottom of the buildings at 0 m. We consider a search radius of 100 m and the number of directions as 16 for SVF calculation (Dirksen et al. 2019). We compute SVF at a resolution of 1 m and later resample it at 30 m resolution.

16.4.4 Solar Radiation

The amount of solar radiation (SR) received by the surfaces in the city influences the LST (Bristow and Campbell 1984; Asadi et al. 2020). Therefore, we compute the SR using the solar radiation tool in ArcGIS Pro 2.9.1. This represents the global radiation, which is the total incoming solar radiation and is calculated for each pixel of DSM. The value of SR was calculated on July 18, 2021 to match the date of acquisition for Landsat-8 image.

16.4.5 Normalized Difference Vegetation Index (NDVI)

NDVI is used to detect bare soil and vegetation (Montandon and Small 2008; Ferreira and Duarte 2019). In a way, it represents the pervious regions in the city. We calculate the NDVI using the Sentinel-2A satellite imagery captured on July 21, 2021, from the United States geological survey (USGS) (https://earthexplorer.usgs.gov/). We chose the image on this date as July and August experience higher temperatures. Moreover, among the images available for this time frame, the selected image had the lowest and most acceptable cloud coverage of less than one per cent. Sentinel-2A image is of 10 m spatial resolution and thus the NDVI obtained is also at 10 m. We calculate NDVI using the near infrared (NIR) and red (R) bands of the image as follows:

$$NDVI = \frac{NIR - R}{NIR + R} \tag{16.3}$$

16.4.6 Normalized Difference Built-Up Index (NDBI)

Zha et al. (2010) defined NDBI to determine urban and built-up areas. It is used to express the intensity of urbanization (Chen et al. 2006). Although we use building density, we consider NDBI as one of the parameters as it helps highlight other urban areas that are not buildings. Moreover, it also explains the development intensity by indicating impervious surfaces. We calculate NDBI using the Sentinel-2A image used for obtaining NDVI. To compute NDBI, we need short wave infrared (SWIR) band, which has a resolution of 20 m and NIR with a resolution of 10 m. Therefore, the SWIR band was resampled to 10 m resolution for the calculation. NDBI is thus calculated as follows:

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \tag{16.4}$$

16.4.7 Frontal Area Index (FAI)

FAI influences the wind flow, thus influencing the LST. It is defined as the area of building walls facing the wind flow in a particular direction (Wong et al. 2010). We compute FAI using the methodology of H. Li et al. (2021) in this paper. The method involves rasterization of the building height and area and computing the FAI at 30 m resolution. The FAI is only calculated for northerly/easterly winds.

16.4.8 Height Variation (HV)

Height Variation (HV) is the variation observed in building heights (1 m pixel) in 30 m pixel. For computing HV, we first transform building heights to a 1 m raster. Then, we aggregate the raster to 30 m with standard deviation of heights in a 30 m pixel using geopandas package in python 3.

16.4.9 Average Height (AH)

Similar to HV, we compute average height (AH) for the pixels by aggregating 1 m height pixels to 30 m, by averaging the heights of 1 m pixels.

16.4.10 Distance to Water

Liege city is situated on the banks of river Meuse. The river divides the city into two parts. Similarly, Brussels has the river Senne that flows through the region. As water bodies have a significant impact on surface temperature (Wu and Zhang 2018), we consider this parameter in our analysis. We obtained the river shapefiles for Liege and Brussels fro, geoportail of Wallonia and UrBIS respectively. We calculate this parameter using Euclidean distance tool to the river shapefiles at 30 m resolution.

16.5 Land Surface Temperature (LST)

We calculate the LST using the Landsat-8 level 1 image captured on 18th July 2021. We choose the image on this date since July and August experience higher temperatures. Moreover, the image on this date had the lowest and most acceptable cloud coverage of less than one per cent. We use the thermal band 10 to compute the LST (in Kelvin (K)) using the following equations (USGS 2019).

$$L_\lambda = M_L Q_{cal} + A_L \tag{16.5}$$

where L_λ = TOA (Top of Atmosphere) spectral radiance (Watts/(m^2 * srad * μm)), M_L = Band-Specific multiplicative rescaling factor from the metadata, A_L = Band-specific additive rescaling factor from the metadata, Q_{cal} = Quantized and calibrated standard product pixel values (DN)

$$T = \frac{K_2}{lnln\left(\frac{K_1}{L_\lambda} + 1\right)} \tag{16.6}$$

where T = TOA brightness temperature (K), K_1 = Band-specific thermal conversion constant from the metadata, K_2 = Band-specific thermal conversion constant from the metadata. We further convert the LST values to degrees Celsius (°C).

16.6 Data Processing

We first generate random points in ArcGIS Pro 2.9.1 over the Brussels capital region and Liege city with 100 m spacing between the points to avoid spatial autocorrelation. We consider 100 m as minimum distance as spatial variability of urban temperatures is around 100 m (Bechtel et al. 2015; Li et al. 2021; Hereher et al. 2022). We then use the "*extract values to point*" tool, to extract values of all the parameters mentioned above to the randomly generated points. We obtain 7500 points in Brussels and 4000 points in Liege, giving us a total of 11,500 points for training and testing the RF regression model.

16.7 RF Regression

RF is a supervised machine learning algorithm proposed by Breiman (2001). The RF regression algorithm combines a large set of regression trees, where dataset is broken down into smaller subsets to predict a response variable by learning decision rules (Breiman et al. 1984). The trees are combined using bootstrap aggregation or bagging, such that each set is run independently, and the outputs are merged to achieve an accurate prediction (Breiman 1996).

The regression begins by selecting n samples of k random observations from a training dataset. Then, individual decision trees are built for each sample. These n trees are run in parallel, and separate outputs are generated. The mean of these outputs results in the final prediction. The random selection in RF regression prevents overfitting (Huynh-Thu and Geurts 2019).

In this study, RF regression is implemented in python through the scikit-learn package. The package includes several parameters that can be tuned for an improved performance. We focus on tuning the number of trees (*ntree*) and number of features randomly sampled at each split (*mtry*).

RF regression is a supervised and straightforward method, which is fast and robust to the noise in the data (Kontschieder et al. 2011; Izquierdo-Verdiguier and Zurita-Milla 2020). We train the RF regressor model using 80% of the data points and test it over remaining 20% of the data points. We validate the RF regression model with *k*-fold cross validation, with *k* set at five. Additionally, we identify the suitable *mtry* and *ntree* based on the lowest value of RMSE (root mean squared error). We also analyse the R-squared value of the relation between observed and predicted LST for both the cities to understand the goodness of fit.

16.8 Simulating Green Roofs

We simulate the green roofs similar to the method proposed by Asadi et al. (2020). Liege has around 20% of flat roofs as computed in the study by Joshi et al. (2020). For Brussels, we identify flat roofs using the 3D model of the region available at UrbIS online. As it will be unrealistic to simulate green roofs on all of these buildings, we consider buildings with area larger than 100 m^2 to be suitable for greening in this study. Figure 16.2 shows the potential roofs in Brussels and Liege. Around 92,333 roofs (out of 256,484) in Brussels are flat, whereas 23,326 roofs (out of 136,170) are flat in Liege.

Green roofs are of mainly of two types: extensive and intensive green roofs. The third type of green roof is a green roof somewhere in between the two main types.

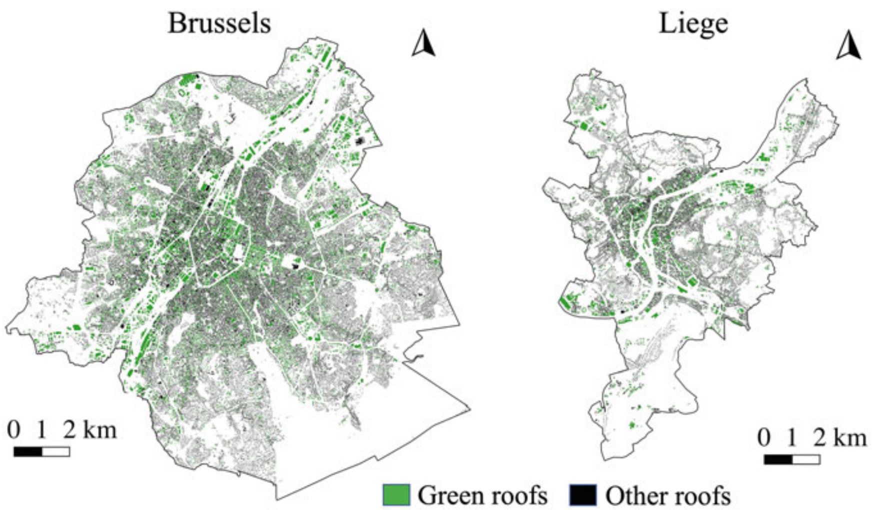


Fig. 16.2 Potential roofs for greening in both the cities

Extensive green roofs have relatively thinner substrate vegetation compared to the intensive ones. Intensive green roofs have a thick substrate layer and dense vegetation, such as rooftop gardens and agriculture (Joshi and Teller 2021). We simulate green roofs by changing the value of NDVI and NDBI for the potential roofs in both cities.

The NDVI values for green surfaces vary from 0.3 to 1 depending upon the intensity of greening. However, for green roofs, NDVI values range from 0.3 to 0.8, considering the range from extensive to intensive green roofs. Further, changing NDBI values also becomes mandatory due to the change in NDVI. We change NDBI values according to the relationship derived between original NDVI and NDBI values obtained from Sentinel-2 for Brussels and Liege. Figure 16.3 represents this correlation.

The relation between NDBI and NDVI is significant, given that the R-squared value is 0.87 and pearson correlation co-efficient is less than 0.05. Based on this relation, Table 16.2 provides the corresponding values of NDBI for each NDVI from 0.3 to 0.8.

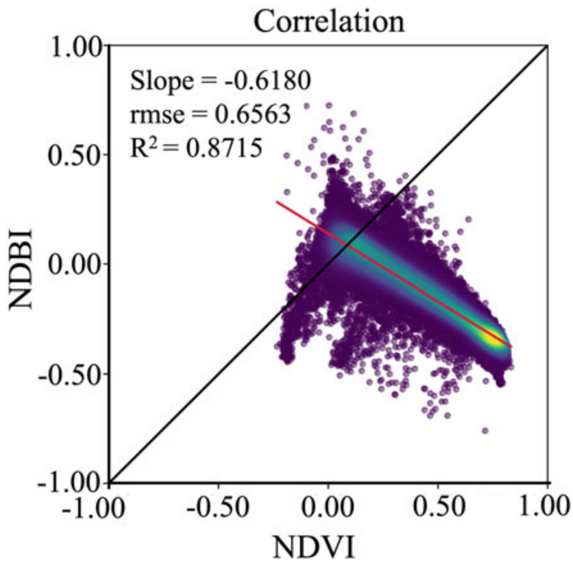


Fig. 16.3 Correlation between NDVI and NDBI values

Table 16.2 Values of NDBI corresponding to NDVI values of green roofs

NDVI (x)	NDBI(y)
0.3	−0.05474677
0.4	−0.11654956
0.5	−0.17835235
0.6	−0.24015514
0.7	−0.30195793
0.8	−0.36376072

Wherever there is a potential roof with 100 m^2, we convert the pixels to NDVI values indicated in the table. We do this by first resampling the 10 m NDVI to 1 m spatial resolution. Thereafter, we change the values of NDVI at the pixels corresponding to building footprints of potential roofs. Later, we aggregate the NDVI with green roofs to 30 m spatial resolution by calculating the average value. Similarly, we convert the NDBI values of potential green roofs to the values in Table 16.1 corresponding to the respective NDVI values. Thus, we generate predictions for six scenarios for six values of NDVI and NDBI in Table 16.1. As only the roof is converted to a "green roof", we keep other building related parameters unchanged. We run the trained model on the newly built NDVI and NDBI along with other variables and predict changes in LST.

16.9 Results

16.9.1 *Model Results and Accuracy*

Figure 16.4 shows the results of RF hyperparameters (*mtry* and *ntree*). We did the optimization based on RMSE. Results indicate that RF hyperparameters affect prediction accuracy only by 0.02 °C. The optimal results are observed at *ntree* = 6000 and *mtry* = 3, with lowest RMSE (1.65 °C).

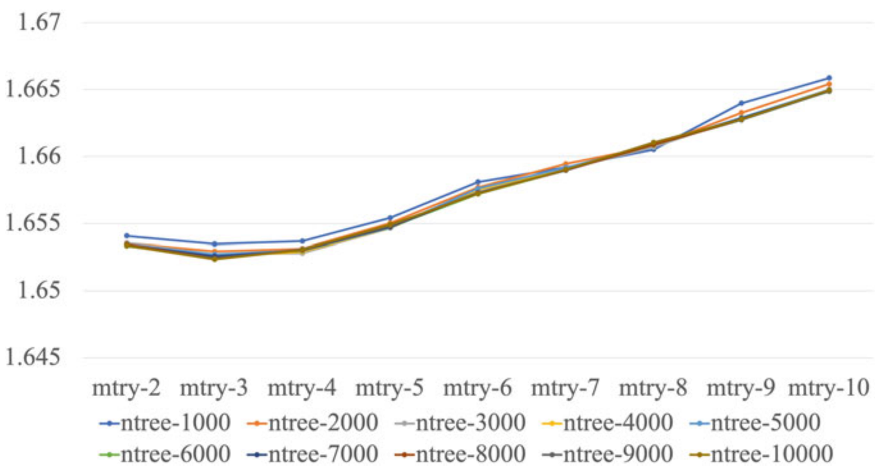


Fig. 16.4 Optimisation of hyperparameters *ntree* and *mtry*

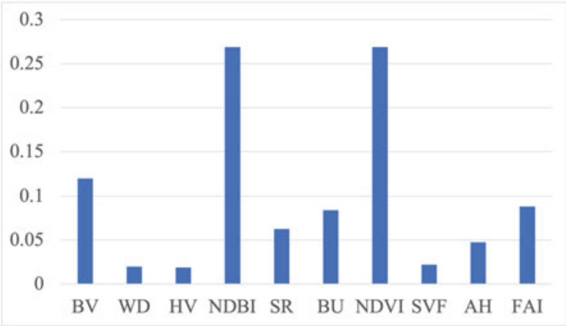


Fig. 16.5 Feature importance of the optimised model

16.9.2 Variable Importance at Optimal Ntree *and* Mtry

Figure 16.5 shows feature importance in an optimized RF regression model. We observe that NDBI and NDVI are the most important parameters, followed by BVI, FAI, AH, SVF, HV and WD.

The model is mainly driven by NDBI, NDVI, BVI and FAI values, which decide the value of LST.

16.9.3 Comparing Predicted and Observed Values of LST

Here, we compare the predicted values of LST with the observed values of LST from Landsat-8 for both the cities. The city of Brussels has observed values of LST ranging from 21 to 42 °C, whereas the city of Liege has the values of LST ranging from 18 to 38 °C (Fig. 16.6). The predicted values for both the cities, however, fall between 21 and 33 °C (Fig. 16.7).

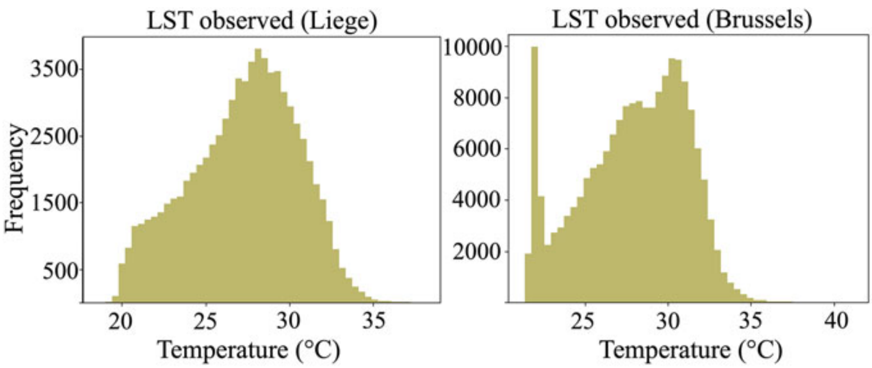


Fig. 16.6 Distribution of observed LST values for Liege and Brussels

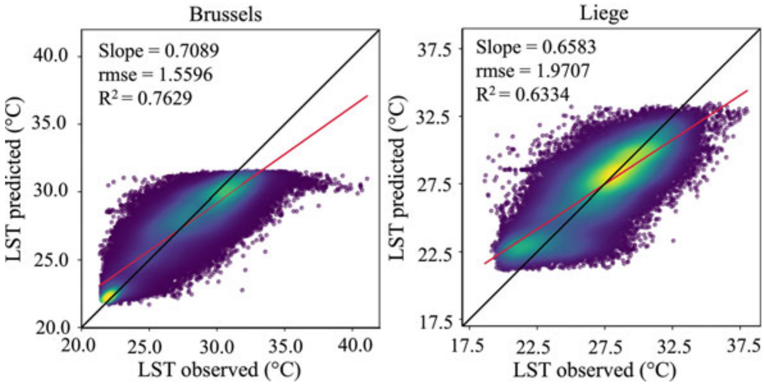


Fig. 16.7 Comparison of predicted vs observed values of LST for Brussels and Liege

We observe that the trained model's R-squared value is 0.76 with an RMSE of 1.55 °C for Brussels. For Liege, the R-squared value is 0.633 with an RMSE of 1.97 °C (Fig. 16.7). We observe that the values between 21 and 33 °C are predicted more accurately compared to the values outside of this range (Fig. 16.7). The reason could be the distribution of data which ranges from 19 to 38 °C, with 80% of the points in the range between 21 and 33 °C. As the model tends to slightly underpredict LST, we compare the effect of green roofs on LST with the predicted LST of our model, to understand the actual impact green roofs can have on LST.

16.9.4 Prediction After Green Roofs

Based on the model, after adding intensive green roofs (NDVI = 0.8) to potential buildings, average LST is shown to be reduced by 0.67 °C and 0.46 °C in Liege, whereas average LST is shown to be reduced by 0.68 and 0.48 °C in the Brussels capital region in building area and entire city respectively. On the other hand, when extensive green roofs (NDVI = 0.3) are added to potential buildings, the average LST can reduce by 0.32 °C and 0.36 °C in Liege, and the average LST can reduce by 0.22 and 0.26 °C in the Brussels capital region in building area and entire city respectively. Figure 16.8 shows the predicted LST for each green roof scenario ranging from NDVI of 0.3–0.8.

Figure 16.9 depicts the distribution of pixels in each class of LST. With increase in NDVI value corresponding to green roofs, there is a decrease in pixels within the range of 32–35 °C for Liege as well as Brussels.

Similarly, Fig. 16.10 depicts the spatial variation in LST with adding intensive and extensive green roofs. As observed in Fig. 16.8, with increase in NDVI values of green roofs, we observe a reduction in pixels with temperature ranges between 29 and 35 °C.

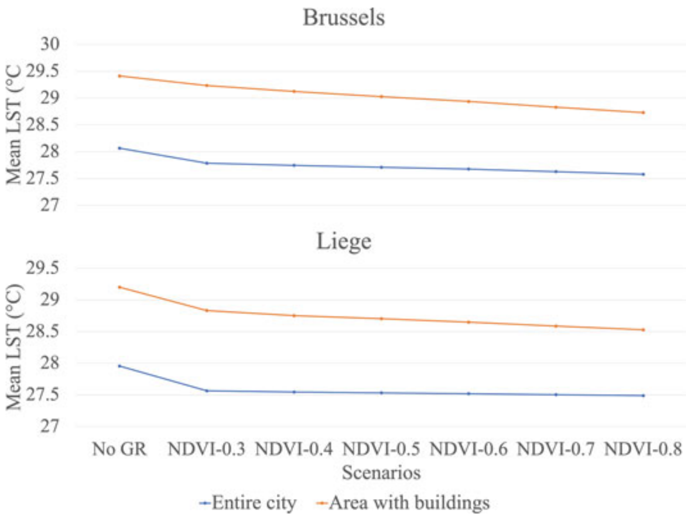


Fig. 16.8 Mean of predicted LST in Liege and Brussels for different scenarios

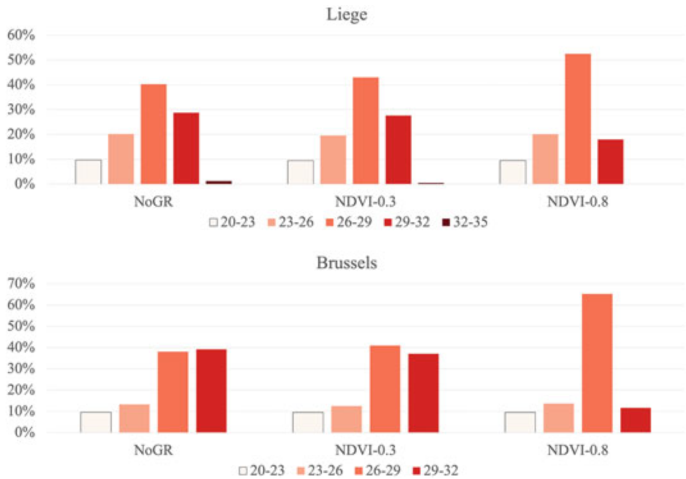


Fig. 16.9 Distribution of proportion of pixels in LST ranges

16.10 Discussion and Conclusions

In this study, we explore the RF regressor model for predicting impact of green roofs on LST. When comparing the observed LST with predicted LST, the model shows a significant goodness of fit. The RF model suggests that green roofs have potential to reduce LST. The benefit of green roofs is higher with intensive green roofs as compared to extensive green roofs.

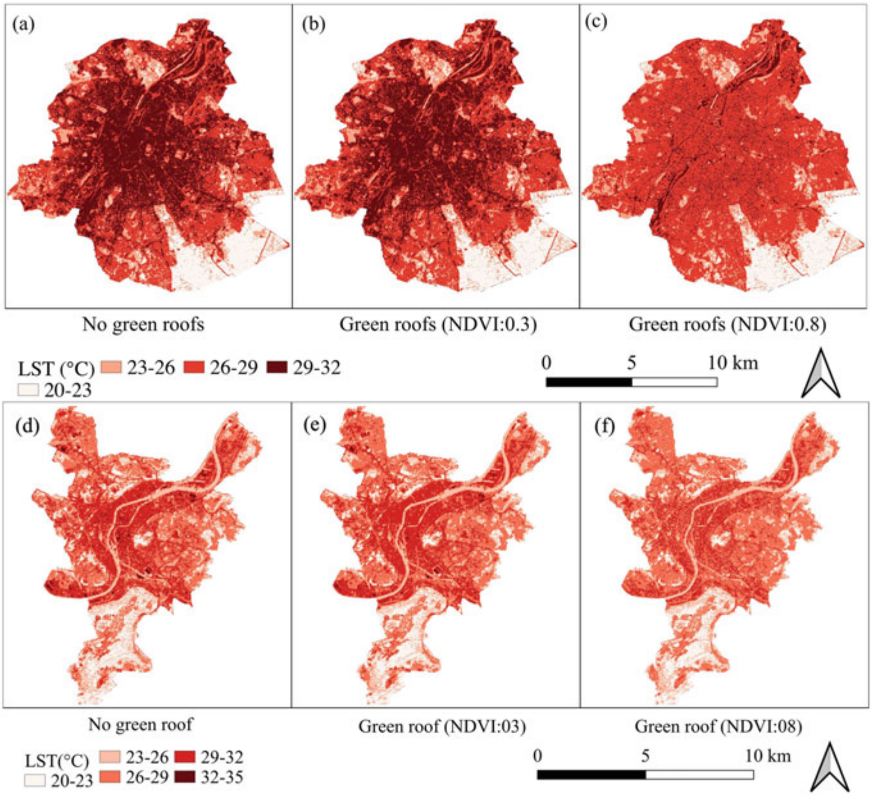


Fig. 16.10 Differences in LST in Brussels and Liege city for different greening scenarios

Although the impact of green roofs on LST on an average seems to be smaller, it is significant in terms of number of pixels where we observe the reduction of temperature (Fig. 16.9). However, small changes in LST may also indicate that when green roofs are placed on existing potential roofs, the impact may not be very significant (at least in case of extensive type of green roofs).

The prediction of LST depends on the training of the RF regressor model. Therefore, we include the data from two major cities in Belgium namely, Brussels and Liege, in order to have sufficient data points for training. However, we observe that the model does not perform well to predict extreme temperatures. To overcome this issue, adding more cities to the dataset can improve prediction accuracy.

Apart from this, the parameter importance of the model implies that NDVI, NDBI and BV govern the predictions. Importance of other parameters such as SVF, WD and HV are relatively low, yet theyare known to be important predictors of the UHI effect (Rodler and Leduc 2019). A reason could be multi-collinearity within the variables. As multi-collinearity does not affect predictions, we consider all the variables to capture maximum variance in the model. For understanding the feature importance, it is important to drop the variables causing multi-collinearity. Further research can combine RF regressor with principal component analysis (PCA) to enhance this analysis.

In this study, we include data points all over the city. However, splitting the training samples into built-up area and non-built up with added cities can improve the prediction accuracy of the model. Additionally, further research can also focus just on analyzing only the built-up area of several cities.

Current model only considers changes in NDVI and NDBI to simulate green roofs. As greening can influence neighborhood areas as well, addition of a neighborhood effect in the model can also increase the prediction accuracy of the model.

Lastly, use of RF regression in prediction of changes in LST after introducing green roofs in a city is a novel and a promising approach, given the proven robustness of RF algorithm in several studies. The model successfully indicates the potential of greening the roofs for reducing the LST in cities. Asadi et al. (2020) performed a similar study using ANN model for Austin, Texas. However, in this study, we introduced two additional parameters influencing wind flow, FAI and HV. We see that FAI influences the model, however, HV has the lowest impact.

References

Adeyeri OE, Akinsanola AA, Ishola KA (2017) Investigating surface urban heat island characteristics over Abuja, Nigeria: relationship between land surface temperature and multiple vegetation indices. Remote Sens Appl 7:57–68. https://doi.org/10.1016/J.RSASE.2017.06.005

Asadi A, Arefi H, Fathipoor H (2020) Simulation of green roofs and their potential mitigating effects on the urban heat island using an artificial neural network: a case study in Austin, Texas. Adv Space Res 66:1846–1862. https://doi.org/10.1016/J.ASR.2020.06.039

Bartesaghi Koc C, Osmond P, Peters A (2018) Evaluating the cooling effects of green infrastructure: a systematic review of methods, indicators and data sources. Sol Energy 166:486–508. https://doi.org/10.1016/J.SOLENER.2018.03.008

Bechtel B, Alexander PJ, Böhner J, et al (2015) Mapping local climate zones for a worldwide database of the form and function of cities. ISPRS Int J Geo-Inform 4:199–219. https://doi.org/10.3390/IJGI4010199

Berardi U, GhaffarianHoseini AH, GhaffarianHoseini A (2014) State-of-the-art analysis of the environmental benefits of green roofs. Appl Energy 115:411–428. https://doi.org/10.1016/J.APENERGY.2013.10.047

Boccalatte A, Fossa M, Gaillard L, Menezo C (2020) Microclimate and urban morphology effects on building energy demand in different European cities. Energy Build 224:110129. https://doi.org/10.1016/j.enbuild.2020.110129

Bowler DE, Buyung-Ali L, Knight TM, Pullin AS (2010) Urban greening to cool towns and cities: a systematic review of the empirical evidence. Landsc Urban Plan 97:147–155. https://doi.org/10.1016/J.LANDURBPLAN.2010.05.006

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. CRC Press

Breiman L (1996) Bagging predictors. Mach Learn 24:123–140. https://doi.org/10.1007/BF00058655

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Bristow KL, Campbell GS (1984) On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agric for Meteorol 31:159–166. https://doi.org/10.1016/0168-1923(84)90017-0

Busato S, Gordon M, Chaudhari M et al (2023) Compositionality, sparsity, spurious heterogeneity, and other data-driven challenges for machine learning algorithms within plant microbiome studies. Curr Opin Plant Biol 71:102326. https://doi.org/10.1016/J.PBI.2022.102326

Chen X-L, Zhao H-M, Li P-X, Yin Z-Y (2006) Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sens Environ 104:133–146. https://doi.org/10.1016/j.rse.2005.11.016

Christis M, Athanassiadis A, Vercalsteren A (2019) Implementation at a city level of circular economy strategies and climate change mitigation—the case of Brussels. J Clean Prod 218:511–520. https://doi.org/10.1016/J.JCLEPRO.2019.01.180

Deilami K, Kamruzzaman M, Liu Y (2018) Urban heat island effect: a systematic review of spatio-temporal factors, data, methods, and mitigation measures. Int J Appl Earth Obs Geoinf 67:30–42. https://doi.org/10.1016/J.JAG.2017.12.009

di Giuseppe E, D'Orazio M (2014) Assessment of the effectiveness of cool and green roofs for the mitigation of the Heat Island effect and for the improvement of thermal comfort in Nearly Zero Energy Building. 58:134–143. https://doi.org/10.1080/00038628.2014.966050

Dirksen M, Ronda RJ, Theeuwes NE, Pagani GA (2019) Sky view factor calculations and its application in urban heat island studies. Urban Clim 30:100498. https://doi.org/10.1016/J.UCLIM.2019.100498

Dong J, Lin M, Zuo J et al (2020) Quantitative study on the cooling effect of green roofs in a high-density urban area—a case study of Xiamen, China. J Clean Prod 255:120152. https://doi.org/10.1016/J.JCLEPRO.2020.120152

El-Zeiny AM, Effat HA (2017) Environmental monitoring of spatiotemporal change in land use/land cover and its impact on land surface temperature in El-Fayoum governorate, Egypt. Remote Sens Appl 8:266–277. https://doi.org/10.1016/J.RSASE.2017.10.003

Ferreira LS, Duarte DHS (2019) Exploring the relationship between urban form, land surface temperature and vegetation indices in a subtropical megacity. Urban Climate 27:105–123. https://doi.org/10.1016/J.UCLIM.2018.11.002

Francis LFM, Jensen MB (2017) Benefits of green roofs: a systematic review of the evidence for three ecosystem services. Urban for Urban Green 28:167–176. https://doi.org/10.1016/J.UFUG.2017.10.015

Govil H, Guha S, Diwan P et al (2020) Analyzing linear relationships of LST with NDVI and MNDISI using various resolution levels of landsat 8 OLI and TIRS data. Adv Intell Syst Comput 1042:171–184. https://doi.org/10.1007/978-981-32-9949-8_13/TABLES/5

Hereher M, Eissa R, Alqasemi A, el Kenawy AM (2022) Assessment of air pollution at Greater Cairo in relation to the spatial variability of surface urban heat island. Environ Sci Pollut Res 29:21412–21425. https://doi.org/10.1007/S11356-021-17383-9/FIGURES/10

Huynh-Thu VA, Geurts P (2019) Unsupervised gene network inference with decision trees and random forests. Methods Mol Biol 1883:195–215. https://doi.org/10.1007/978-1-4939-8882-2_8/COVER

Izquierdo-Verdiguier E, Zurita-Milla R (2020) An evaluation of guided regularized random forest for classification and regression tasks in remote sensing. Int J Appl Earth Obs Geoinf 88:102051. https://doi.org/10.1016/J.JAG.2020.102051

Jamei E, Chau HW, Seyedmahmoudian M, Stojcevski A (2021) Review on the cooling potential of green roofs in different climates. Sci Total Environ 791:148407. https://doi.org/10.1016/J.SCITOTENV.2021.148407

Jato-Espino D, Manchado C, Roldán-Valcarce A, Moscardó V (2022) ArcUHI: a GIS add-in for automated modelling of the urban heat Island effect through machine learning. Urban Clim 44:101203. https://doi.org/10.1016/J.UCLIM.2022.101203

Joshi MY, Rodler A, Musy M et al (2022) Identifying urban morphological archetypes for microclimate studies using a clustering approach. Build Environ 224:109574. https://doi.org/10.1016/J.BUILDENV.2022.109574

Joshi MY, Teller J (2021) Urban integration of green roofs: current challenges and perspectives. Sustainability 13:12378. https://doi.org/10.3390/SU132212378

Joshi MY, Selmi W, Binard M, et al (2020) Potential for urban greening with green roofs: a way towards smart cities. https://doi.org/10.5194/isprs-annals-VI-4-W2-2020-87-2020

Kim SW, Brown RD (2021) Urban heat island (UHI) variations within a city boundary: a systematic literature review. Renew Sustain Energy Rev 148:111256. https://doi.org/10.1016/j.rser.2021.111256

Kleerekoper L, van Esch M, Salcedo TB (2012) How to make a city climate-proof, addressing the urban heat island effect. Resour Conserv Recycl 64:30–38. https://doi.org/10.1016/j.resconrec.2011.06.004

Kokalj Ž, Somrak M (2019) Why not a single image? Combining visualizations to facilitate fieldwork and on-screen mapping. Remote Sens 11:747. https://doi.org/10.3390/RS11070747

Kontschieder P, Bulò SR, Bischof H, Pelillo M (2011) Structured class-labels in random forests for semantic image labelling. Proc IEEE Int Conf Comp Vis 2190–2197. https://doi.org/10.1109/ICCV.2011.6126496

Li H, Liu Y, Zhang H et al (2021) Urban morphology in China: dataset development and spatial pattern characterization. Sustain Cities Soc 71:102981. https://doi.org/10.1016/j.scs.2021.102981

Li D, Bou-Zeid E, Oppenheimer M (2014) The effectiveness of cool and green roofs as urban heat island mitigation strategies. Environ Res Lett 9. https://doi.org/10.1088/1748-9326/9/5/055002

Lin M, Dong J, Jones L et al (2021) Modeling green roofs' cooling effect in high-density urban areas based on law of diminishing marginal utility of the cooling efficiency: a case study of Xiamen Island, China. J Clean Prod 316:128277. https://doi.org/10.1016/J.JCLEPRO.2021.128277

Lyu F, Wang S, Han SY, et al (2022) An integrated cyberGIS and machine learning framework for fine-scale prediction of Urban Heat Island using satellite remote sensing and urban sensor network data. Urban Inform 1:1–15. https://doi.org/10.1007/S44212-022-00002-4

Matsuki K, Kuperman V, van Dyke JA (2016) The Random Forests statistical technique: an examination of its value for the study of reading. Sci Stud Read 20:20–33. https://doi.org/10.1080/10888438.2015.1107073/SUPPL_FILE/HSSR_A_1107073_SM1111.DOCX

Mirzaei PA (2015) Recent challenges in modeling of urban heat island. Sustain Cities Soc 19:200–206. https://doi.org/10.1016/J.SCS.2015.04.001

Montandon L, Small E (2008) The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. Remote Sens Environ 112:1835–1845. https://doi.org/10.1016/j.rse.2007.09.007

Razzaghmanesh M, Beecham S, Salemi T (2016) The role of green roofs in mitigating urban heat Island effects in the metropolitan area of Adelaide, South Australia. Urban for Urban Green 15:89–102. https://doi.org/10.1016/J.UFUG.2015.11.013

Rodler A, Leduc T (2019) Local climate zone approach on local and micro scales: dividing the urban open space. Urban Clim 28:100457. https://doi.org/10.1016/j.uclim.2019.100457

Roth M, Oke TR, Emery WJ (1989) Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology. Int J Remote Sens 10(11):1699–1720

Santamouris M (2013) Energy and climate in the urban built environment. Routledge, London

Stewart ID, Oke TR (2012) Local climate zones for urban temperature studies. Bull Am Meteorol Soc 93:1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1

USGS (2019) Landsat 8 (L8) Data Users Handbook. USGS, Sioux Falls, South Dakota

Wang X, Li H, Sodoudi S (2022) The effectiveness of cool and green roofs in mitigating urban heat island and improving human thermal comfort. Build Environ 217:109082. https://doi.org/10.1016/J.BUILDENV.2022.109082

Wang B, Geoffroy S, Bonhomme M (2021) Urban form study for wind potential development. 49:76–91. https://doi.org/10.1177/2399808321994449

Wong MS, Nichol JE, To PH, Wang J (2010) A simple method for designation of urban ventilation corridors and its application to urban heat island analysis. Build Environ 45:1880–1889. https://doi.org/10.1016/j.buildenv.2010.02.019

Wu Z, Zhang Y (2018) Spatial variation of urban thermal environment and its relation to green space patterns: implication to sustainable landscape planning. Sustainability 10:2249. https://doi.org/10.3390/SU10072249

Yang J, Bou-Zeid E (2019) Scale dependence of the benefits and efficiency of green and cool roofs. Landsc Urban Plan 185:127–140. https://doi.org/10.1016/J.LANDURBPLAN.2019.02.004

Zakšek K, Oštir K, Kokalj Ž (2011) Sky-view factor as a relief visualization technique. Remote Sensing 3:398–415. https://doi.org/10.3390/RS3020398

Zha Y, Gao J, Ni S (2010) Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. 24:583–594. https://doi.org/10.1080/01431160304987

Chapter 17
A Framework to Probe Uncertainties in Urban Cellular Automata Modelling Using a Novel Framework of Multilevel Density Approach: A Case Study for Wallonia Region, Belgium

Anasua Chakraborty, Ahmed Mustafa, Hichem Omrani, and Jacques Teller

Abstract Urban expansion models are widely used to understand, analyze and predict any peculiar scenario based on input probabilities. Modelling and uncertainty are concomitant, and can occur due to reasons ranging from–discrepancies in input variables, unpredictable model parameters, spatio-temporal variability between observations, or malfunction in linking model variables under two different spatio-temporal scenarios. However, uncertainties often occur because of the interplay of model elements, structures, and the quality of data sources employed; as input parameters influence the behavior of cellular automaton (CA) models. Our study aims to address these uncertainties. While most studies consider neighborhood effects, timestep and spatial resolution, our study uniquely focuses on the susceptibility of multi density classes and varying cell size on uncertainty. Hence this chapter offers a theoretical elucidation of the concepts, sources, and strategies for managing uncertainty under various criteria as well as an algorithm for enumerating the model's accuracy for Wallonia, Belgium.

Keywords Uncertainty analysis · Urban CA modelling · Urban densification

---

A. Chakraborty (✉) · J. Teller
LEMA Research Group, Urban & Environmental Engineering Dept, University of Liège, Liège, Belgium
e-mail: A.Chakraborty@uliege.be

J. Teller
e-mail: Teller@uliege.be

A. Mustafa
The New School, Urban Systems Lab, New York, NY, USA

H. Omrani
Urban Development and Mobility, Luxembourg Institute of Socio-Economic Research, Esch-Sur-Alzette, Luxembourg
e-mail: hichem.omrani@liser.lu

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0_17

17.1 Introduction

Over the last few decades, pressures on land availability brought on by urbanization has become a central issue, particularly for developing countries. The enormous land requirement to serve a rapidly expanding urban population has presented significant impediments to economic prosperity, social inclusion, and environmental protection (Angel et al. 2021; Jiang et al. 2022). Various modelling methods have been developed in recent years to handle various urban challenges which differ in approach and underlying principles (Li and Gong 2016). For example, several statistical and geospatial urban models have been designed to analyze the relationship between driving forces and urban land change, and predict its future development, including Agent-based models (ABM) (Zhang et al. 2010), logistic regression (Mustafa et al. 2018a), cellular automata (Almeida et al. 2003; García et al. 2013), and Conversion of Land use and its effects (CLUE) model (Verburg et al. 2014). Urban planners and decision-makers can better comprehend the environmental and socioeconomic elements that encourage urbanization trends thanks to the application of these urban development models (Batisani and Yarnal 2009).

In general, CA is viewed as a bottom-up urban model from which emergent patterns of land use change are produced from 'simple' transition rules, which is in line with the core principles of complexity science; the interactions of simple subsystems lead to the formation of complex systems (Lu et al. 2019). Uncertainty can be measured through probability distribution and involves risks. Uncertainty with a known events of possible outcomes and quantifiable probability involves risk which are known as objective risks while the one whose outcomes are purely reliant on human judgements are known as subjective risk (Loucks et al. 2005). Model's uncertainty can be a causative effect of various type of errors. The error can occur due to model parameters or input variables. Vardoulakis et al. (2002) has used different dispersion model as he suggests that it is impossible to assess the uncertainty of a model based on discrete values of input variables. On the other hand, Gar-On and Li (n.d.) in their work have experimented with errors in acquiring and applying GIS data sources that are used as a input variables while simulating numerous CA models, thus helping to achieve realistic and accurate results.

Urban CA models, like any other urban models, are subject to numerous sources of uncertainty which are difficult to disentangle. Uncertainty might be viewed as a measure of how much we distrust the concepts and abstractions we use to represent the real world. An archetypal urban CA model is made up of four parts: transition rules, neighborhood configuration, simulation time (time step), and stochastic perturbation (Yeh and Li 2006). Due to the complex characteristics among each component, urban CA results may be sensitive to variations in parameter settings and adopted approaches (García et al. 2011; Li et al. 2014). These inaccuracies will spread during CA simulation and have an impact on the results of the simulation. In order to do this, the effects of source errors and error propagation on simulation outcomes must be assessed. While, Uncertainty is crucial because it enhances the model's accuracy and account the variability in inputs, and allows folding up the uncertainty of input into the set ofoutput values (*Sensitivity and Uncertainty–Center for Systems Reliability* 2022). Conversely, sensitivity analysis examines the model's resilience, and evaluates the influence of a model's assumptions, thus concluding that inputs with greater impact on output sensitivity provide the effect of interactions between input components. Hence, uncertainty analysis seeks to quantify the ambiguity in a model's result. The purpose of sensitivity analysis is to determine how variation in input values corresponds to variance in output measurements. It is done by altering one or more input variables, and measuring the effect on output measurements (*Sensitivity and Uncertainty–Center for Systems Reliability* 2022).

The simplicity with which CA models may currently be integrated and programmed in raster-based geographic information system (GIS) systems (Kocabas and Dragicevic 2006), as well as with other methodologies like agent-based or multi-criteria assessment (Batty 2016; Wu 2016), is one reason why they are receiving more and more focus. Simple rules can be used in CA models to produce complicated patterns (Wolfram 2002). By configuring fundamental CA model components including cell states, cell size, neighborhood size and type, transition rules, and temporal increments, it is feasible to properly reflect spatial complexity and the dynamics of urban development change (Torrens and O'Sullivan 2022; White and Engelen 2000; Yen and Li 2016). Among these, spatial extent is frequently linked to a specific research case, preventing it from having the universal property of spatial scale sensitivity during CA-based land use change simulation. There haven't been enough studies to date that have carried out a systematic investigation of the consequences of changing the CA model design's constituent parts during the calibration process (Wu et al. 2019). While most papers deal with several kind of model-based or input-oriented uncertainties and their sensitivity towards cellular automata modelling, we have attempted to present in our work a theoretical prototype of evaluation taking into consideration a multi-level density approach. It involves changing input variables and observing their error propagation, but under different urban built-up density classes, thus helping the original data to remain untouched and resulting in realistic model simulation outputs.

In the remainder of the paper, we provide background information on various types of uncertainty analysis based on cell size, transition rules and density classes. In Sect. 17.3, we explain the methodology for analyzing uncertainty based on urban CA models in a univariate method and how it can be implemented for Wallonia region, Belgium. Subsequently, in Sect. 17.4 we present results and discuss the influence of cell size, transition rules and density class on the model output. In Sect. 17.5, we conclude by providing a list of future tasks that can be implemented in order to investigate sensitivity and uncertainty analysis using urban CA model.

17.2 Background

When applied to real cities, urban CA models are prone to errors and uncertainty. Uncertainty in geospatial data might be unavoidable when developing a realistic simulation. However, differing combinations of neighborhood size, type, and spatial resolution can have an impact on simulation outcomes. By perturbing spatial variables and analyzing the error terms in the simulation results, one may easily scrutinize error propagation in urban simulation. The variability in the model output is quantified by uncertainty analysis (UA), and the distribution of this uncertainty across the model input factors is examined by sensitivity analysis (SA) (Crosetto and Tarantola 2010; Saltelli et al. 1999). Sensitivity analysis examines the link between input and output information in a model and pinpoints the sources of variation that affect model outputs. Uncertainty can come from a variety of sources, including errors and approximations in the measurement of the input data, parameter values, model structure, and model solution techniques. Input errors and model flaws both spread during the simulation process in CA simulation.

While prior research has offered an examination of CA model behavior in relation to modifying the model components, several recent studies have focused on the topic of errors and uncertainties associated with CA models. Variations in cell size, transition rule, and multi-level density classes have been explored in the current literature. In order to examine how uncertainty affects simulation outcomes for the Ourthe River basin in Wallonia, Belgium, Mustafa et al. (2014) employed a stochastic component. This component addressed how randomness propagates in urban growth models. The model was validated using cell-to-cell allocation and validated using landscape matrices. The findings showed that as stochastic perturbation size grows, the accuracy of the model declines, starting with extremely tiny perturbations. Feng and Liu (2013) introduced a unique geographic-based CA model that incorporates sensitivity analysis techniques to optimize transition rules that were first constructed using logistic regression. Menard and Marceau also explored how spatial scale affected the results of the CA model in relation to the spatial resolution and a typical CA neighborhood layout. They expanded on their study even further and introduced VecGCA (Moreno et al. 2008), a new object-based geographical cellular automata model. This model was a versatile and effective tool for simulating changes in land use and cover as well as other spatiotemporal occurrences that suggested object geometry alterations. VecGCA ensured a more accurate portrayal of geographic space (as well as growth of the items constituting it) by being independent of cell size, neighborhood arrangement, and landscape configuration by incorporating a dynamic neighborhood.

17.2.1 Approaches in Evaluating Uncertainty Based on Scale Effects: A Multiscale Approach

Scale may be understood as a continuum across which things, patterns, and processes can be seen and connected (Marceau 2014). In relative space, scale is a variable that is inextricably tied to the spatial entities, patterns, forms, as well as functions, processes, and rates under consideration. In absolute space, scale refers to a practical, standard system used to divide geographical space into an operational spatial unit (Samat 2006). Marceau (2014), on the other hand, provides a concept of scale that is generally accepted when discussing urban CA models and relates it to the absolute and relative depiction of space.

The scale problem has been addressed using a variety of methods. It runs the gamut of from straightforward scale analysis methods like geographic variance, local variation, and texture analysis, to more intricate methods like semi-variograms and fractals (Chen and Henebry 2009; García-Álvarez et al. 2019a, b; Young et al. 2021; Zhang et al. 2019). According to earlier research by Kok and Veldkamp (2001) on the impact of modifying scale in a LUCC model for Central America, increasing the resolution from 15*15 km to 75*75 km enhanced the model's explanatory power (r^2) but had no discernible impact on the explanatory factors. In their study of the behavior of several urban development rules at various cell sizes in the well-known SLEUTH (Clarke 2008) model, Jantz and Goetz (2007) came to the conclusion that the resolution of the cells had a significant role in the performance of the model, and that some urban growth rules produced significantly greater growth at coarser resolutions than at finer ones. Although their results are highly unique to the SLEUTH model, the conclusion is that neighborhood effects for urban land, which are essential to all CA models, may vary non-linearly across scales.

17.2.2 Scenario Description and Impacts of Uncertainty on Transition Rules

Transition rules in CA model are essentially designed such that a cell's future state depends on both its current state and the state of its surroundings (Liu 2008; Stevens et al. 2007; Ward et al. 2003). In order to characterize land-use transition potential of a cell, many driving mechanisms for urban landuse change have been discovered, mapped, and implemented in GIS (Liu et al. 2008; Wu 2010). These driving forces or factors may be divided into three major groups of variables: socioeconomic, geophysical, and accessibility and have been listed in Table 17.1 (Chakraborty et al. 2022; Mustafa et al. 2018a, b, c).

Many transition rules, including those based on Markov chains (Kamusoko et al. 2009), neural networks (Li and Yeh 2010b), genetic algorithms (Shan et al. 2008), ant colony optimization (Liu et al. 2010), cuckoo search algorithms (Cao et al. 2015), particle swarm optimization (Feng et al. 2018), and data mining (Li and Yeh 2010a), have been developed based on various intrinsic principles of land use conversion. Although a CA-Markov only models land use changes using a constant time step (Pontius and Malanson 2007), one significant advantage is that it simultaneously predicts the trajectory of land use change among various categorical states (Li et al. 2016). It also simulates spatio-temporal dynamic changes (Jenerette and Wu 2001)

Table 17.1 Varied metrics of senstivity analysis used in the past

S.No	Publication	Metric	Aspect	Study area	KEY strength
1	Jantz and Goetz (2007)	Scale Sensitivity	Scale	Washington, DC–Baltimore	Scale as an integral issue during each phase of a modelling effort
2	Kocabas and Dragicevic (2006)	Neighborhood sensitivity	Neighborhood	San Diego region, USA	Exploratory method of sensitivity analysis that can be used in a gen-eral context to examine and assess CA model errors and uncertainties
3	Feng and Liu (2013)	Transition sensitivity	Transition rules	Shanghai, China	Model the multiple land-use changes and spatial ecological processes using SA optimisation algorithm
4	Ménard and Marceau (2016)	Scale sensitivity	Scale	Maskoutains region (Quebec, Canada)	Finer exploration of cell size sensitivity by using a stochastic geographic cellular automata (GCA) model
5	Moreno et al. (2008)	Neighborhood sensitivity	Neighborhood	Quebec, Canada	Size sensitivity by allowing the representation of space as a collec-tion of geographic objects by im-plementing a novel VectorGCA

(continued)

Table 17.1 (continued)

S.No	Publication	Metric	Aspect	Study area	KEY strength
6	Samat (2006)	Scale sensitivity	Scale	Seberang Perai region, Penang State, Malaysia	GIS based CA model using multi-criteria evaluation (MCE) suitability index map
7	Ward et al. (2003)	Transition sensitivity	Transition rules	Queensland, Australia	Address strategic planning and management issues by integrating regional and local scale models
8	Wu (2010)	Transition sensitivity	Transition rules	Guangzhou, South China	Stochastic CA model for rural–urban land conversion using Monte Carlo process

that attempt to improve the simulation accuracy as much as possible. Mustafa et al. (2021) forecasted future land use changes in New York City by introducing a multi-objective Markov Chain Monte Carlo (MO-MCMC) that considered multiple allocation objectives. They concluded that the model can bring analytical and simulation approaches into the planning process, and simulate trade-offs between different LULC visions.

17.3 Materials and Methods

17.3.1 Study Area

Wallonia, located in the southern part of Belgium, accounts for 55% of total area of the country with a coverage of 16,844 km^2 (Mustafa et al. 2018a), as shown in Fig. 17.1. It predominantly consists of the French speaking area and is mostly characterized by peri-urban areas. Despite its large areal extent, Wallonia comprises of only one third of the total Belgian population. The important urban centers of Wallonia are Liege, Mons, Namur, and Charleroi. The urban density of the region has a peculiar pattern of sparse settlements in most of the outskirts with populated inhabitants mostly in the metropolitan centers (Li et al. 2015).

This area mostly stretches from Eastern part (e.g., Liege) along the industrial zones to the west in Mons. Wallonia, with its urban development design, makes it

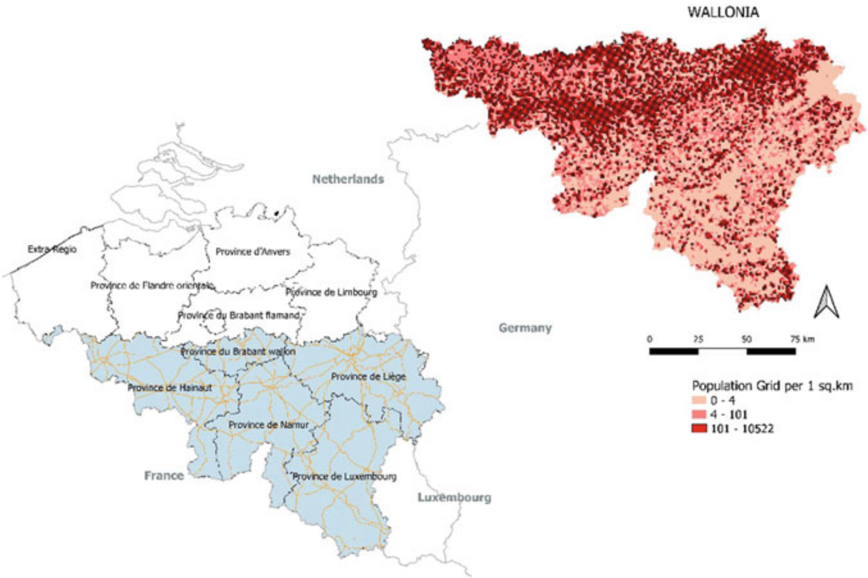


Fig. 17.1 Wallonia region (southern Belgium) with population grid per sq.km

useful for studying dispersed urbanization patterns. It is also an area that has been less researched in the CA urban growth literature. Hence this region was selected for modelling urban densification, along with its associated uncertainties.

17.3.2 Conceptual Framework

In Fig. 17.2, we show our conceptual method which includes our diverse datasets and modelling structures. This includes our previous study that used a novel approach. In this work (Mustafa et al. 2018b), a temporal Monte Carlo method (TMC) was proposed to study the uncertainty over time for Liege, Wallonia, with a study being conducted between year 1990 and 2010 for a short term simulation This is different from other models that use randomness in order to study uncertainity. Further, this model was subsequently extended for a longer term till year 2100.

Our literature study showed that the basic necessities for a standard CA model involved implementing neighborhood effect, transition rules and cell size for simulating development in future. García-Álvarezet al. (2019a, b) and Gar-On Yeh usually concentrate on technical uncertainty, which takes into account usual errors related to spatial scale or raster data classification. While it is evident from literature that uncertainty can be due to input variables, it is important to note that driving factors are also crucial. A range of socioeconomic, physical, accessibility and environmental variables have been considered, which when calibrated using a logit model, can

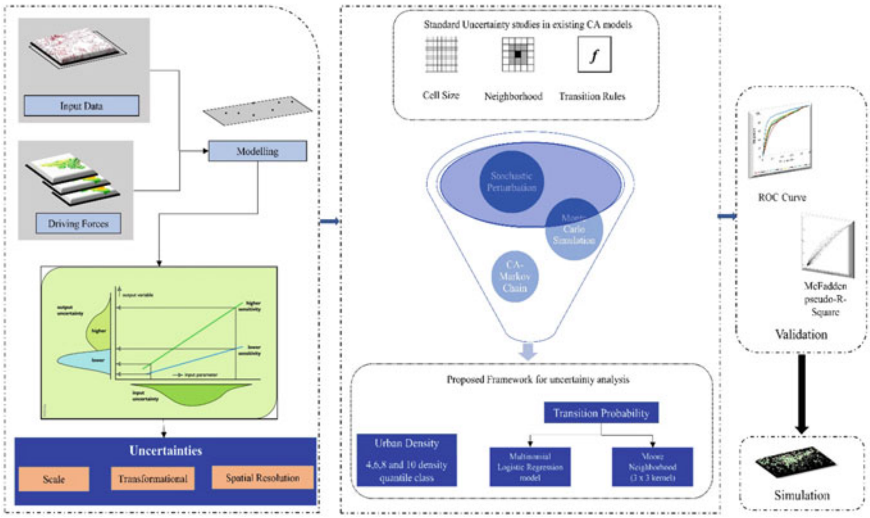


Fig. 17.2 Conceptual methodological framework for uncertainty analysis. *Source* partly adapted from (Loucks et al. 2005)

help provide better validation. Therefore, while most studies have the tendency of following a traditional input-based or model-based study, it is imperative to consider uncertainties of spatial allocation as well (Mustafa et al. 2018b).

By observing the land use pattern at time t and time t + 1, the calibration process aims to derive the coefficients or parameter values for CA transition rules. Using a multinomial logistic model, the likelihood of development may be determined when there are numerous factors. The probability is calculated by comparing land use changes in CA over a longer time period than one cycle. The estimated transition probability of a randomly chosen cell is compared with a uniform random number within a dynamic range at each time-step. The methodologies of Wu et al. as well as Mustafa et al. (2018a), were used to determine the transition probabilities (2002). Mathematically, the transition probability P for cell *i* at time-step *t* can be calculated as follows:

$$Pi^t = \left(Pi^d\right) \times \left(Pi^n\right)^t \times con(.)$$

where $\left(Pi^t\right)$ is the urbanization probability based on driving forces, $(Pi^n)^t$ is the neighborhood interaction, and $con(.)$ Is restrictive constant for land use change.

Hence, as a part of our previous research, we attempted to carry out the study on a vector based Cadastral data which was further converted into raster format at a scale of 100*100 m representing different density classes for urban built up development. Since most of the uncertainty studies on CA model involves urban expansion as their phenomenon, It is crucial to take into account the uncertainty that occurs as a result of change in density classes of different combinations across time and space. This is the main objective for our hypothesis.

17.4 Observation and Assessments

In this section, we highlight the different strategies that have been suggested till now, with ultimate goal for dealing with uncertainty analysis using CA. The purpose of CA models is to truly portray urban and land-use patterns by reflecting their changes over time. Considering that complexity is a state between order and chaos, and that creating complexity necessitates the introduction of the proper amount of randomness in the models, it becomes vital to assess the best strategy for doing so. Different strategies are consequently required for a thorough analysis of uncertainty because there is no single method that can analyze all sorts of uncertainty. For this reason, the paper explored the uncertainty analysis in a multivariate approach.

Comparing the model's predictions to actual data in order to determine its level of uncertainty is the process of validating a cellular automata (CA) model of urban expansion. This might involve contrasting the simulated urban land-use patterns, population density, and other spatial aspects of the city with actual data in the context of urban growth modelling. Comparing the simulated land-use patterns to satellite imagery or other remote sensing data is a typical method for evaluating a CA model of urban expansion (Chaudhuri and Clarke 2014) This may be done by visualizing a comparison between the satellite picture and the generated land-use map. Another strategy is to contrast crucial model statistics with actual data, such as the percentage of urban land.

Comparing the model's predictions against those of other models or data sources is another method of model validation. An official census or other sources of demographic information can be used to compare the model's estimates of population density, for instance. The model's performance may also be evaluated in relation to other urban growth models that employ other techniques or premises. By repeatedly running the model with various initial conditions or parameter values and comparing the outcomes, the degree of uncertainty may be assessed. This can assist pinpoint the model's sources of uncertainty and show how sensitive the model is to changes in the inputs.

Although transition rules and transition probability are essential to CA modelling, CA models are also influenced by neighborhood configuration and scaling behavior, as well as scaling behaviors of transition probability (Dahal and Chow 2015; Feng and Tong 2018). As a result, even though transition rules and transition probability are crucial to CA modelling, they are not the only factors that have an impact on simulation results (Gao et al. 2020). However, there is a relationship between model performance and the amount of change across the simulation time. As discussed previously, typical CA models include a variety of intrinsic model uncertainties that are connected to the neighborhood, cell size, computation time, transition rules, and model parameters (Gar-On Yeh and Li n.d.).

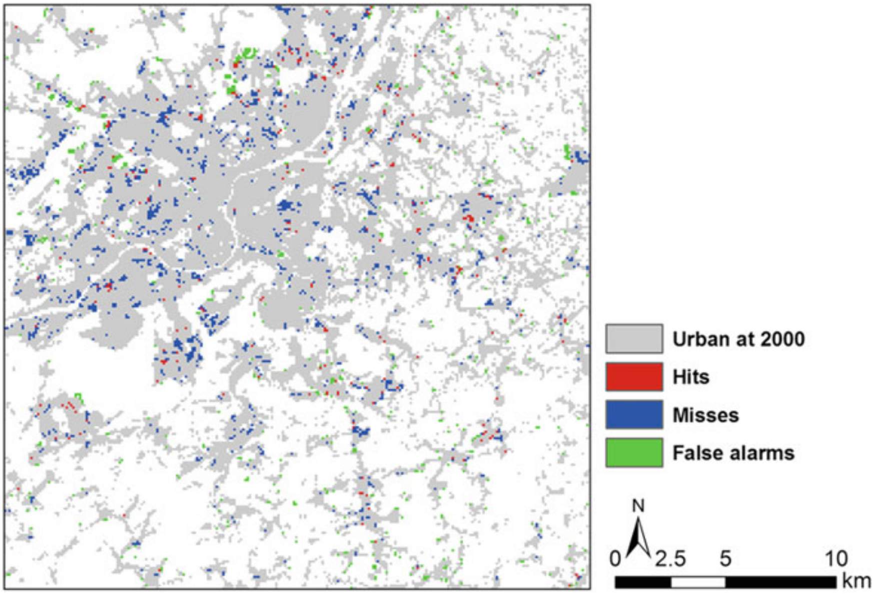


Fig. 17.3 The hits, misses, false alarms during the validation time interval for Liege metropolitan

When applied to a local context, such as Wallonia, Mustafa et al.(2018b) proposed a Time Monte Carlo (TMC) method to introduce randomness in land use change models with the aim of modelling spatial allocation uncertainty. By analysing (a) the hits (H) which indicate that the areas of expansion on the observed map were simulated as expansion, it evaluated the allocation performance; (b) misses (M) that indicate that the areas of expansion on the observed map were simulated as no changes; (c) false alarms (FA), indicating that the no-changes in the observed map were simulated as expansion; and (d) correct rejections (CR), indicating that the areas of no-change in the observed map were simulated correctly (Fig. 17.3).

Mustafa et al. (2018b) further transformed the transition probability of each cell by comparing it with the largest available probability at each time-step, following Wu's (2010) study. Mathematically it can be written as:

$$Pi'^{t} = Pi^{t}\exp\left[\delta(1 - Pi^{t}/\max\left(P^{t}\right)\right]$$

where Pi'^{t} is the updated transition probability for cell i at time-step t, Pi^{t} is the original probability, δ is a dispersion term, and $\max(P^{t})$ finds the maximum transition probability at time-step t. The dispersion term in a cellular automata (CA) simulation is used to measure the degree of variation or spread in the density of the system at different levels. The density of a CA system refers to the number of active or "on" cells at a given time, and the dispersion term is used to measure how this density varies across different spatial scales or levels. This information can then be used to influence the behavior of the simulation, for example by adjusting the rules for

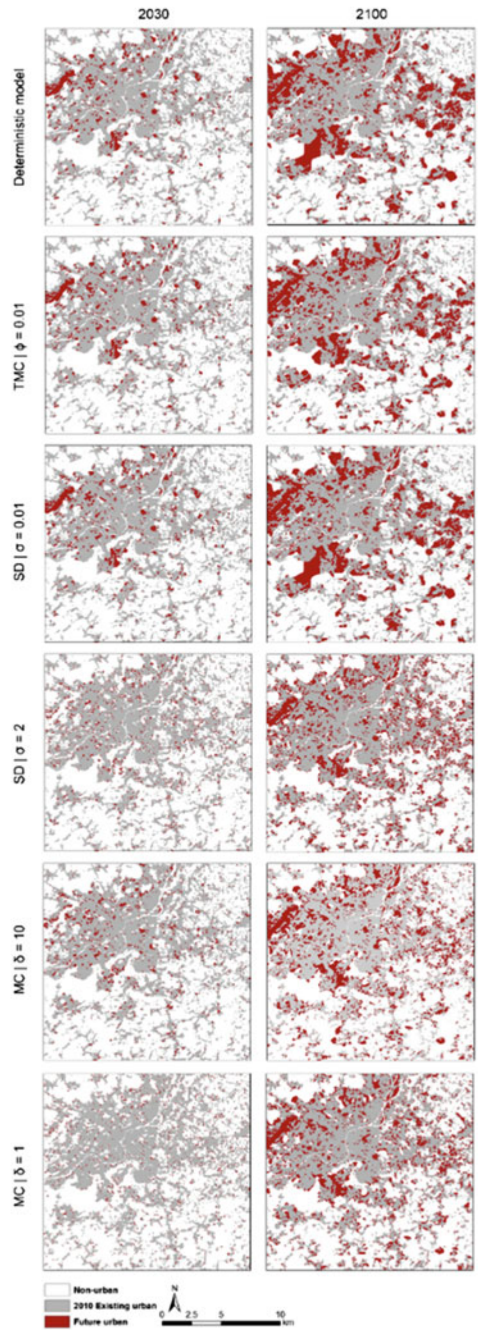


Fig. 17.4 2030 and 2100 simulations for different configurations for Liege metropolitan

cell activation or interaction based on the observed density variations. The skewed probability curve's form is controlled by the dispersion parameter.

To further investigate the model's performance, different values of δ have been used to generate future urban patterns for 2030 and 2100 (Fig. 17.4). At early time-steps (like 2030), the proposed TMC model generates simulations that are comparable to the outcomes from deterministic models, with minor changes. However, the model generates results different from a deterministic-equation based simulation model as it predicts further into the future.

17.5 Conclusion

In this study, we demonstrate the uncertainty with regards to CA model for analyzing urban complexities. This uncertainty is mostly a result of different cell size, neighborhood dynamics and transition rules. Which is why the uncertainty can be model input-based or model parameter based. Because our work involves urban densification, it is crucial to understand the effect of density classes on uncertainty. Hence, we proposed a theoretical framework explaining the establishment of uncertainty in models through past works carried out for Liege. We employed the TMC model to evaluate uncertainty, and forecasted urban development for 2030 and 2100. In our future research, we will extend the current study that shows uncertainty resulting from applying a CA model based on multi-level urban density to a larger area. Further, as a part of our future work, we intended to apply the model uncertainty for simulating residential densification. Finally, a simplified CA-MCMC (Cellular Automata Monte Carlo) simulation could be used for other parts of Belgium, for example the Brussels capital region, and Flemish and Wallon Brabant. Also, unlike the usual variables that are commonly used for such studies (e.g., cell neighborhood and cell size), novel variables like cell density was applied in our study. Thus, going forward this research framework will facilitate in adding a different dimension to uncertainty studies in urban modeling, specifically in studying densification while achieving valid simulation results.

Funding: This research was funded by the INTER program and co-funded by the Fond National de la Recherche, Luxembourg (FNR) and the Fund for Scientific Research-FNRS, Belgium (F.R.S—FNRS), T.0233.20,—'Sustainable Residential Densification' project (SusDens, 2020–2023).

References

Almeida CM, Monteiro AMV, Câmara G, Soares-Filho BS, Cerqueira GC, Araújo WL, Pantuzzo AE (2003) Modeling the urban evolution of land use transitions using cellular automata and logistic regression. Int Geosci Remote Sens Symp (IGARSS) 3:1564–1566. https://doi.org/10.1109/IGARSS.2003.1294176

Angel S, Lamson-Hall P, Blei A, Shingade S, Kumar S (2021) Densify and expand: a global analysis of recent urban growth. Sustainability 13(7):3835. https://doi.org/10.3390/SU13073835

Batisani N, Yarnal B (2009) Uncertainty awareness in urban sprawl simulations: lessons from a small US metropolitan region. Land Use Policy 26(2):178–185. https://doi.org/10.1016/J.LANDUSEPOL.2008.01.013

Batty M (2016) Urban Evolution on the desktop: simulation with the use of extended cellular automata. Environ Plan A Econ Space 30(11):1943–1967. https://doi.org/10.1068/A301943

Cao M, Tang G, Shen Q, Wang Y (2015) A new discovery of transition rules for cellular automata by using cuckoo search algorithm. Int J Geogr Inf Sci 29(5):806–824. https://doi.org/10.1080/13658816.2014.999245

Chakraborty A, Omrani H, Teller J (2022) Modelling the drivers of urban densification to evaluate built-up areas extension: A data-modelling solution towards zero net land take. In: Lecture notes in computer science (Including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 13376 LNCS. pp 260–270. https://doi.org/10.1007/978-3-031-10450-3_21/COVER

Chaudhuri G, Clarke KC (2014) Temporal accuracy in urban growth forecasting: a study using the SLEUTH model. Trans GIS 18(2):302–320. https://doi.org/10.1111/TGIS.12047

Chen W, Henebry GM (2009) Change of spatial information under rescaling: a case study using multi-resolution image series. ISPRS J Photogramm Remote Sens 64(6):592–597. https://doi.org/10.1016/J.ISPRSJPRS.2009.05.002

Clarke KC (2008) A decade of cellular urban modeling with SLEUTH: unresolved issues and problems, Ch. 3. In: Planning support systems for cities and regions, pp 47–60. http://bbs.geog.ucsb.edu

Crosetto M, Tarantola S (2010) Uncertainty and sensitivity analysis: tools for GIS-based model implementation. Int J Geogr Inf Sci 15(5):415–437. https://doi.org/10.1080/13658810110053125

Dahal KR, Chow TE (2015) Characterization of neighborhood sensitivity of an irregular cellular automata model of urban growth. Int J Geogr Inf Sci 29(3):475–497. https://doi.org/10.1080/13658816.2014.987779

Feng Y, Wang J, Tong X, Liu Y, Lei Z, Gao C, Chen S (2018) The effect of observation scale on urban growth simulation using particle swarm optimization-based CA models. Sustainability 10:4002. https://doi.org/10.3390/SU10114002

Feng Y, Liu Y (2013) A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing. Int J Geogr Inf Sci 27(3):449–466. https://doi.org/10.1080/13658816.2012.695377

Feng Y, Tong X (2018) Dynamic land use change simulation using cellular automata with spatially nonstationary transition rules. Gisci Remote Sens 55(5):678–698. https://doi.org/10.1080/15481603.2018.1426262

Gao C, Feng Y, Tong X, Lei Z, Chen S, Zhai S (2020) Modeling urban growth using spatially heterogeneous cellular automata models: comparison of spatial lag, spatial error and GWR. Comput Environ Urban Syst 81:101459. https://doi.org/10.1016/J.COMPENVURBSYS.2020.101459

García AM, Santé I, Crecente R, Miranda D (2011) An analysis of the effect of the stochastic component of urban cellular automata models. Comput Environ Urban Syst 35(4):289–296. https://doi.org/10.1016/J.COMPENVURBSYS.2010.11.001

García AM, Santé I, Boullón M, Crecente R (2013) Calibration of an urban cellular automaton model by using statistical techniques and a genetic algorithm. Application to a small urban settlement of NW Spain. Int J Geograph Inform Sci 27(8):1593–1611. https://doi.org/10.1080/13658816.2012.762454

García-Álvarez D, Camacho Olmedo MT, Paegelow M (2019a) Sensitivity of a common Land Use Cover Change (LUCC) model to the Minimum Mapping Unit (MMU) and Minimum Mapping Width (MMW) of input maps. Comput Environ Urban Syst 78:101389. https://doi.org/10.1016/J.COMPENVURBSYS.2019.101389

García-Álvarez D, van Delden H, Teresa Camacho Olmedo M, Paegelow M, García-Álvarez D, Camacho Olmedo MT, van Delden H, Paegelow Geode M, Paegelow M (2019b). Uncertainty challenge in geospatial analysis: an approximation from the land use cover change modelling perspective. 289–314. https://doi.org/10.1007/978-3-030-04750-4_15

Gar-On Yeh A, Li X (n.d.) Error propagation and model uncertainties of cellular automata in urban simulation with GIS

Jantz CA, Goetz SJ (2007) Analysis of scale dependencies in an urban land-use-change model. Int J Geograph Inform Sci 19(2):217–241. https://doi.org/10.1080/13658810410001713425

Jenerette GD, Wu J (2001) Analysis and simulation of land-use change in the central Arizona—Phoenix region, USA. Landscape Ecol 16(7):611–626. https://doi.org/10.1023/A:1013170528551

Jiang H, Guo H, Sun Z, Xing Q, Zhang H, Ma Y, Li S (2022) Projections of urban built-up area expansion and urbanization sustainability in China's cities through 2030. J Clean Prod 367:133086. https://doi.org/10.1016/J.JCLEPRO.2022.133086

Kamusoko C, Aniya M, Adi B, Manjoro M (2009) Rural sustainability under threat in Zimbabwe—simulation of future land use/cover changes in the Bindura district based on the Markov-cellular automata model. Appl Geogr 29(3):435–447. https://doi.org/10.1016/J.APGEOG.2008.10.002

Kocabas V, Dragicevic S (2006) Assessing cellular automata model behaviour using a sensitivity analysis approach. Comput Environ Urban Syst 30(6):921–953. https://doi.org/10.1016/J.COMPENVURBSYS.2006.01.001

Kok K, Veldkamp A (2001) Evaluating impact of spatial scales on land use pattern analysis in Central America. Agr Ecosyst Environ 85(1–3):205–221. https://doi.org/10.1016/S0167-8809(01)00185-2

Li X, Gong P (2016) Urban growth models: progress and perspective. Sci Bull 61(21):1637–1650. https://doi.org/10.1007/S11434-016-1111-1

Li X, Yeh AGO (2010a) Data mining of cellular automata's transition rules. Int J Geogr Inf Sci 18(8):723–744. https://doi.org/10.1080/13658810410001705325

Li X, Yeh AGO (2010b) Neural-network-based cellular automata for simulating multiple land use changes using GIS. Int J Geogr Inf Sci 16(4):323–343. https://doi.org/10.1080/13658810210137004

Li S, Colson V, Lejeune P, Speybroeck N, Vanwambeke SO (2015) Agent-based modelling of the spatial pattern of leisure visitation in forests: a case study in Wallonia, south Belgium. Environ Model Softw 71:111–125. https://doi.org/10.1016/J.ENVSOFT.2015.06.001

Li X, Yu L, Sohl T, Clinton N, Li W, Zhu Z, Liu X, Gong P (2016) A cellular automata downscaling based 1 km global land use datasets (2010–2100). Sci Bull 61(21):1651–1661. https://doi.org/10.1007/S11434-016-1148-1

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. 28(7):1317–1335. https://doi.org/10.1080/13658816.2014.883079

Liu X, Li X, Shi X, Wu S, Liu T (2008) Simulating complex urban development using kernel-based non-linear cellular automata. Ecol Model 211(1–2):169–181. https://doi.org/10.1016/J.ECOLMODEL.2007.08.024

Liu X, Li X, Liu L, He J, Ai B (2010) A bottom-up approach to discover transition rules of cellular automata using ant intelligence. Int J Geogr Inf Sci 22(11–12):1247–1269. https://doi.org/10.1080/13658810701757510

Liu Y (2008) Modelling urban development with geographical information systems and cellular automata. CRC PRESS. https://www.routledge.com/Modelling-Urban-Development-with-Geographical-Information-Systems-and-Cellular/Liu/p/book/9780367577438

Loucks DP, van Beek E, Stedinger JR, Dijkman JPM, Villars MT (2005) Water resources systems planning and management an introduction to methods, models and applications

Lu Y, Laffan S, Pettit C, Cao M (2019) Land use change simulation and analysis using a vector cellular automata (CA) model: a case study of Ipswich City, Queensland, Australia. Environ Plan B Urban Anal City Sci 47(9):1605–1621. https://doi.org/10.1177/2399808319830971

Marceau DJ (2014) The scale issue in the social and natural sciences. Can J Remote Sens 25(4):347–356. https://doi.org/10.1080/07038992.1999.10874734

Ménard A, Marceau DJ (2016) Exploration of spatial scale sensitivity in geographic cellular automata. Environ Plan B Urban Analyt City Sci 32(5):693–714. https://doi.org/10.1068/B31163

Moreno N, Ménard A, Marceau DJ (2008) VecGCA: a vector-based geographic cellular automata model allowing geometric transformations of objects. Environ Plan B Urban Analyt City Sci 35(4):647–665. https://doi.org/10.1068/B33093

Mustafa A, Saadi I, Cools M, Teller J (2014) Measuring the effect of stochastic perturbation component in cellular automata urban growth model. Procedia Environ Sci 22:156–168. https://doi.org/10.1016/J.PROENV.2014.11.016

Mustafa A, Saadi I, Cools M, Teller J (2018a) Understanding urban development types and drivers in Wallonia: a multi-density approach. Int J Bus Intell Data Mining 13(1–3):309–330. https://doi.org/10.1504/IJBIDM.2018.088434

Mustafa A, van Rompaey A, Cools M, Saadi I, Teller J (2018c) Addressing the determinants of built-up expansion and densification processes at the regional scale. Urban Stud 55(15):3279–3298. https://doi.org/10.1177/0042098017749176

Mustafa A, Ebaid A, Omrani H, McPhearson T (2021) A multi-objective Markov Chain Monte Carlo cellular automata model: simulating multi-density urban expansion in NYC. Comput Environ Urban Syst 87:101602. https://doi.org/10.1016/J.COMPENVURBSYS.2021.101602

Mustafa A, Saadi I, Cools M, Teller J (2018b) A Time Monte Carlo method for addressing uncertainty in land-use change models. 32(11), 2317–2333. https://doi.org/10.1080/13658816.2018.1503275

Pontius GR, Malanson J (2007) Comparison of the structure and accuracy of two land change models. Int J Geogr Inf Sci 19(2):243–265. https://doi.org/10.1080/13658810410001713434

Pontlus RG (2000) Quantification error versus location error in comparison of categorical maps. Photogramm Eng Remote Sens 66(8):1011–1016

Saltelli A, Tarantola S, Chan KPS (1999) A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 41(1):39–56. https://doi.org/10.1080/00401706.1999.10485594

Samat N (2006) Characterizing the scale sensitivity of the cellular automata simulated urban growth: a case study of the Seberang Perai Region, Penang State, Malaysia. Comput Environ Urban Syst 30(6):905–920. https://doi.org/10.1016/J.COMPENVURBSYS.2005.11.002

Sensitivity and Uncertainty–Center for Systems Reliability. (n.d.). Available at 22 Oct 2022 https://www.sandia.gov/csr/center-for-systems-reliability/capabilities/sensitivity-and-uncertainty/

Shan J, Alkheder S, Wang J (2008) Genetic algorithms for the calibration of cellular automata urban growth modeling. Photogramm Eng Remote Sens 74(10):1267–1277. https://doi.org/10.14358/PERS.74.10.1267

Stevens D, Dragicevic S, Rothley K (2007) iCity: a GIS–CA modelling tool for urban planning and decision making. Environ Model Softw 22(6):761–773. https://doi.org/10.1016/J.ENVSOFT.2006.02.004

Torrens PM, O'Sullivan D (2022) Cities, cells, and complexity: developing a research agenda for urban geocomputation. In: Abrahart RJ, Carlisle BH (eds) International conference on geocomputation. Available at 24 Oct 2022 http://www.geocomputation.org/2000/GC044/Gc044.htm

Vardoulakis S, Fisher BEA, Gonzalez-Flesca N, Pericleous K (2002) Model sensitivity and uncertainty analysis using roadside air quality measurements. Atmos Environ 36(13):2121–2134. https://doi.org/10.1016/S1352-2310(02)00201-7

Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura SSA (2014) Modeling the spatial dynamics of regional land use: the CLUE-S model. Environ Manag 30(3):391–405. https://doi.org/10.1007/S00267-002-2630-X

Ward DP, Murray AT, Phinn SR (2003) Integrating spatial optimization and cellular automata for evaluating urban change. Ann Reg Sci 37(1):131–148. https://doi.org/10.1007/S001680200113

White R, Engelen G (2000) High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Comput Environ Urban Syst 24(5):383–400. https://doi.org/10.1016/S0198-9715(00)00012-0

Wolfram S (2002) Stephen Wolfram: A new kind of science. Wolfram Media Inc. https://www.wolframscience.com/nks/

Wu F (2010) Calibration of stochastic cellular automata: the application to rural-urban land conversions. Int J Geogr Inf Sci 16(8):795–818. https://doi.org/10.1080/13658810210157769

Wu F (2016) An experiment on the generic polycentricity of urban growth in a cellular automatic city. Environ Plan B Urban Analyt City Sci 25(5):731–752. https://doi.org/10.1068/B250731

Wu H, Li Z, Clarke KC, Shi W, Fang L, Lin A, Zhou J (2019) Examining the sensitivity of spatial scale in cellular automata Markov chain simulation of land use change. Int J Geogr Inf Sci 33(5):1040–1061. https://doi.org/10.1080/13658816.2019.1568441

Yeh AGO, Li X (2006) Errors and uncertainties in urban cellular automata. Comput Environ Urban Syst 30(1):10–28. https://doi.org/10.1016/J.COMPENVURBSYS.2004.05.007

Yen AGO, Li X (2016) A constrained CA model for the simulation and planning of sustainable urban forms by using GIS environment and planning b: urban analytics and city science 28(5):733–753. https://doi.org/10.1068/B2740

Young SG, Datta J, Kar B, Huang X, Williamson MD, Tullis JA, Cothren J (2021) Challenges and limitations of geospatial data and analyses in the context of COVID-19. In: Mapping COVID-19 in space and time. Springer, Cham, pp 137–167. https://doi.org/10.1007/978-3-030-72808-3_8

Zhang B, Xu G, Jiao L, Liu J, Dong T, Li Z, Liu X, Liu Y (2019) The scale effects of the spatial autocorrelation measurement: aggregation level and spatial resolution. Int J Geograph Inform Sci 33(5):945–966. https://doi.org/10.1080/13658816.2018.1564316

Zhang H, Zeng Y, Bian L, Yu X (2010) Modelling urban expansion using a multi agent-based model in the city of Changsha. J Geograph Sci 20(4):540–556. https://doi.org/10.1007/S11442-010-0540-Z

Index

R. Goodspeed et al. (eds.), *Intelligence for Future Cities*, The Urban Book Series,
https://doi.org/10.1007/978-3-031-31746-0