sustainability

Special Issue Reprint

Sustainable City Planning and Development

Transport and Land Use

Edited by
Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

mdpi.com/journal/sustainability

MDPI

Sustainable City Planning and Development: Transport and Land Use

Sustainable City Planning and Development: Transport and Land Use

Editors

Linchuan Yang
Yuanyuan Guo
Yaoming Zhou
Wenxiang Li
Jixiang Liu

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Editors
Linchuan Yang
Urban and Rural Planning
Southwest Jiaotong University
Chengdu
China

Yuanyuan Guo
Spatial Planning and Landscape Design
Guangzhou University
Guangzhou
China

Yaoming Zhou
Industrial Engineering and Management
Shanghai Jiao Tong University
Shanghai
China

Wenxiang Li
Transportation System Engineering
University of Shanghai for Science and Technology
Shanghai
China

Jixiang Liu
Urban Planning
Xiamen University
Xiamen
China

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Sustainability* (ISSN 2071-1050) (available at: www.mdpi.com/journal/sustainability/special_issues/sustainable_cities_transport).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-1318-6 (Hbk)
ISBN 978-3-7258-1317-9 (PDF)
doi.org/10.3390/books978-3-7258-1317-9

Contents

About the Editors

Linchuan Yang

Prof. Linchuan Yang is a professor and deputy head of the Department of Urban and Rural Planning at the School of Architecture, Southwest Jiaotong University, located in Chengdu, China. He obtained both Ph.D. and M.Phil. degrees from The University of Hong Kong and a bachelor's degree in engineering and a bachelor's degree in science from Xiamen University, China. His research interests include transport and land use, transport planning, travel behavior, transport geography, transport and city planning, and the built environment.

He actively contributes to the academic community through his editorial roles. He serves as an editorial board member of *Transportation Research Part D* and as a guest editor for esteemed journals such as *Transportation Research Part D*, *Travel Behaviour and Society*, and the *Journal of Housing and the Built Environment*. Additionally, he assumes the responsibility of associate editor for *Heliyon* and *Frontiers in Public Health*.

His research endeavors have received support from prominent institutions such as the Ministry of Science and Technology of China, the National Natural Science Foundation of China, the Science & Technology Department of Sichuan Province, the Philosophy and Social Science Planning Office of Sichuan Province, and the Peking University-Lincoln Institute. He has served as a peer reviewer for over 90 SCI/SSCI journals.

Yuanyuan Guo

Dr. Yuanyuan Guo is an associate professor in the Department of Spatial Planning and Landscape Design of the College of Architecture and Urban Planning, Guangzhou University, also working as the off-campus mentor of the Department of Urban and Rural Planning at Tianjin University.

He is a member of the editorial board of several scientific journals, including *Digital Transportation and Safety* and *Urbanism and Architecture*. He acts as a reviewer for over 30 international and Chinese scientific journals in the fields of transportation, urban planning, and geography. In addition, he is a member of the International Association for China Planning (IACP), The American Association of Geographers (AAG), and the Urban Planning Society of China (UPSC).

His research areas include travel behavior, urban transport and land use, shared mobility, GIS, and spatial analysis. He has participated in several funded research projects, including the National Natural Science Foundation of China and the Tianjin Applied Basic Research Youth Project. He has over 50 published works in international scientific journals, book chapters, conference proceedings, and other scientific outlets.

Yaoming Zhou

Dr. Yaoming Zhou received a B.Eng. degree in Mechatronics from Zhejiang University, China in 2014 and a Ph.D. degree in Industrial Engineering from The University of Hong Kong in 2018. From 2018 to 2019, he was a senior algorithm engineer with Alibaba Cainiao Logistics. Since 2019, he has been an associate professor at the Department of Industrial Engineering and Management, Shanghai Jiao Tong University, Shanghai, China.

His research interests include the modeling and analysis of transportation systems; the integration of operations research, data analytics, and artificial intelligence; and their application to transportation. He has published over 30 works in international journals, such as *Transportation Research Part A/B/E* and *IEEE Transactions on Intelligent Transportation Systems*.

Wenxiang Li

Dr. Wenxiang Li is an associate professor and deputy director of the Institute of Regional Planning and Transportation Systems at the University of Shanghai for Science and Technology. He obtained his bachelor's degree in Transportation Engineering from Harbin Institute of Technology and his Ph.D. in Transportation Planning and Management from Tongji University. He was also a visiting scholar at the University of Washington. He is a youth committee member of the Small-Town Planning Committee of the Urban Planning Society of China and serves as a guest editor for several international journals.

His recent research focuses on low-carbon transportation planning, spatiotemporal big data mining, shared mobility management, and travel service optimization. He has led seven research projects funded by the National Natural Science Foundation of China, the Science and Technology Commission of Shanghai Municipality, the Shanghai Social Science Planning Office, and the Shanghai Education Development Foundation. He has published over 60 academic papers in peer-reviewed journals, with more than 40 as the first or corresponding author and over 30 indexed by SCI/SSCI. He also holds two authorized invention patents and has published one academic monograph.

Jixiang Liu

Dr. Jixiang Liu is an assistant professor in the Department of Urban Planning of the School of Architecture and Civil Engineering at Xiamen University. He obtained his bachelor's, master's, and doctorate degrees successively from Central South University, Peking University, and The University of Hong Kong. His research interests include, among others, travel behavior, urban mobility, urban vibrancy, machine learning, and big data.

He has published more than 30 articles in international peer-reviewed journals. He has received funding support from diverse organizations including the National Natural Scientific Foundation of China (NSFC). Also, he has served as an anonymous reviewer for over 20 international journals.

Preface

The sustainable development of transport and land use represents a pivotal aspect of sustainable city planning and development. Historically, the evolution of transport systems has profoundly influenced urban land use patterns. For example, in North America, from the development of streetcar suburbs in the late 19th century to the rise of car-dependent sprawl in the mid-20th century, transport innovations have shaped where people live, work, and play. Similarly, land use decisions, such as zoning regulations and development policies, have affected the layout of transport infrastructure and the accessibility of different areas in the city. Today, as cities grapple with congestion, pollution, and the excess emissions of greenhouse gases, understanding the relationship between transport and land use is more critical than ever. City planners and policymakers recognize the need to promote sustainable transport modes, such as active travel modes, while also fostering compact, mixed-use developments that minimize the need for long-distance trips and decrease the number of motorized trips.

The relationship between transport and land use is complex and dynamic, with each influencing the other in significant ways. For example, efficient transport systems increase spatial accessibility to different areas in the city, thereby influencing land values. Locations with easy access to transport hubs, such as rail and metro stations, tend to have higher land values and are often developed more intensively. Moreover, transport networks shape land use. The layout of roads and transit routes can influence the density and layout of urban areas. Compact, mixed-use developments tend to be more conducive to efficient public transport systems, while sprawling suburbs often rely more heavily on cars. In addition, land use patterns can influence people's mode choice for transport. Dense urban areas with mixed-use zoning often encourage walking, cycling, and the use of public transport, while low-density suburban areas may necessitate car travel for most daily activities. Furthermore, transport infrastructure can drive land use planning, for example, through transit-oriented development (TOD) initiatives. TOD aims to promote mixed-use, pedestrian-friendly developments around transit stations, encouraging people to live, work, and play within walking distance of transit stations. In recent years, the discourse around the interactions between transport and land use has gained prominence, highlighting its significance.

In essence, fostering sustainable city planning and development requires a holistic understanding of the complex interactions between transport and land use. In other words, to promote sustainable city planning and development, systematic and thorough assessments and a deeper comprehension of the interactions between transport and land use are needed, especially for today's cities.

Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li, and Jixiang Liu
Editors

sustainability

Editorial

Sustainable City Planning and Development: Transport and Land Use

Linchuan Yang

Department of Urban and Rural Planning, School of Architecture, Southwest Jiaotong University, Chengdu 611756, China; yanglc0125@swjtu.edu.cn

1. Background

The Sustainable Development Goals (SDGs) are a set of 17 global goals established by the United Nations' *Transforming our World: the 2030 Agenda for Sustainable Development* in 2015 as a universal call to action to address various social, economic, and environmental challenges facing the world. The SDGs, such as No Poverty (SDG 1), Zero Hunger (SDG 2), and Good Health and Well-being (SDG 3), are designed to tackle a wide range of global challenges, promote sustainable development in many facets, and improve quality of life for people around the world while emphasizing a commitment to leaving no one behind. These goals provide a shared blueprint for countries, businesses, and individuals to work toward a more sustainable and equitable future by 2030.

Referred to as "Sustainable Cities and Communities", SDG 11 seeks to address the extensive challenges associated with urbanization and promote the development of inclusive, safe, resilient, and sustainable cities and communities. At its core, SDG 11 recognizes the unprecedented trend of urbanization, with over half of the world's population now residing in cities. This goal acknowledges the opportunities and complexities arising from this demographic shift, emphasizing the need for cities and human settlements to prioritize environmental sustainability, social inclusivity, and economic prosperity. In essence, SDG 11 envisages cities and communities as dynamic hubs of innovation, opportunity, and inclusivity. Through addressing various dimensions of urban/community development, the goal encourages us to create cities and communities that provide a high quality of life for all residents. Notably, achieving SDG 11 requires collaborative efforts from governments, local communities, businesses, and international organizations, emphasizing the importance of shared responsibility in building sustainable cities and communities for the future.

A prerequisite for achieving SDG 11 is the sustainable development of transport (spatial interaction) and land use (spatial development) [1]. Transport and land use are intricately linked components that jointly shape the physical, social, and economic fabric of cities, communities, and human settlements [2,3]. It is widely accepted that transport affects land use through accessibility, while land use affects transport through activities (Figure 1). More specifically, on the one hand, land use generates travel/activity demand and profoundly affects residents' trip production, attraction, distribution, and choice of mode of transport, thus influencing their travel patterns. On the other hand, transport changes the accessibility of land and the demand for urban development. This relationship plays a pivotal role in determining the efficiency, sustainability, and livability of cities, communities, and human settlements. Evidently, it is very complex and requires careful consideration and examination.

For an extended period of time, in numerous countries, including China, there has been a notable disconnection (or, at least, a semi-disconnection) between urban transport planning and land-use planning. Addressing this deficiency is key to resolving various interconnected urban challenges, such as traffic congestion and urban sprawl. The crucial solution lies in the strategic integration of urban transport and land-use planning. Therefore,

check for updates

Citation: Yang, L. Sustainable City Planning and Development: Transport and Land Use. *Sustainability* **2023**, *15*, 15804. https://doi.org/10.3390/su152215804

Received: 4 November 2023
Accepted: 8 November 2023
Published: 10 November 2023

in recent years, the topic of coordinating transport and land use has entered the public discourse, and its significance has been greatly emphasized.

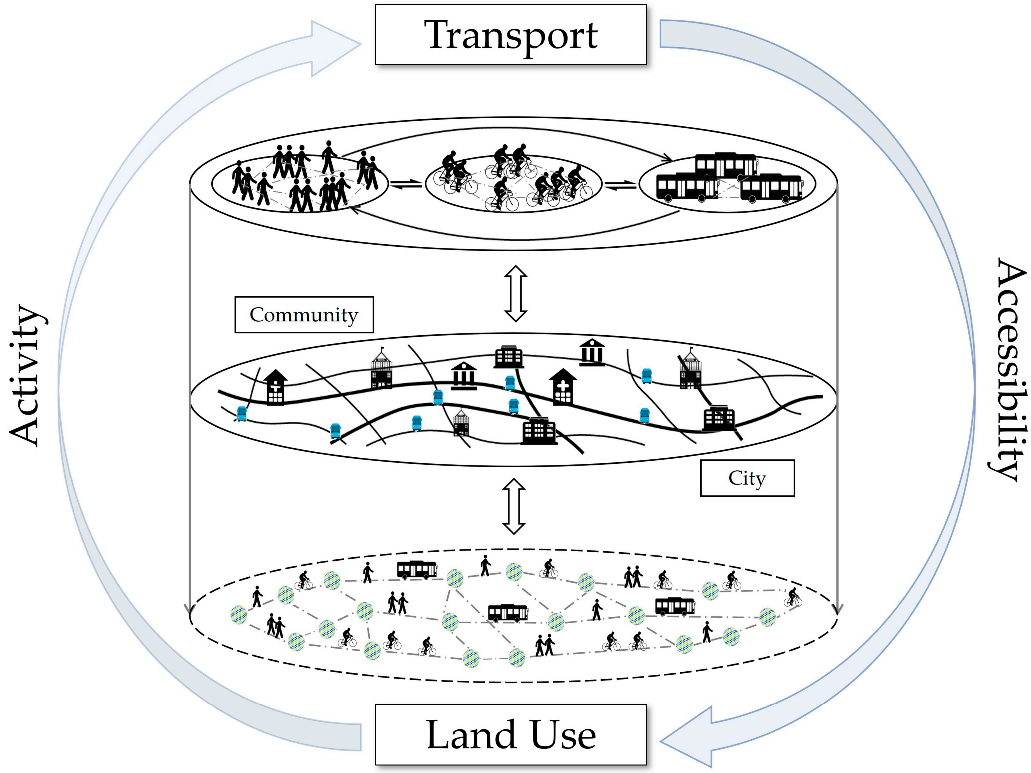


Figure 1. The interaction loop between transport and land use.

Systematic and rigorous evaluations and an enriched understanding of the interactions between transport and land use are urgently required to inform decision making and guide urban/transport planning, management, and development [4]. This is crucial to achieving transport–land-use integration and meeting SDG 11. The multifaceted connections between transport and land use need to be investigated, such as their historical evolution, contemporary challenges, and the critical role they play in the broader context of sustainable development, as well as approaches to the integration and coordinated development of transport and land use.

2. Overview of this Special Issue

This Special Issue opened in the Section "Sustainable Urban and Rural Development" of *Sustainability* in September 2021 and closed in July 2023. Its Guest Editors are five early career researchers specializing in Urban and Rural Planning or Transportation Planning and Management, including Dr. Linchuan Yang (Professor of Urban and Rural Planning at Southwest Jiaotong University), Dr. Yuanyuan Guo (Associate Professor of Urban and Rural Planning at Tianjin University), Dr. Yaoming Zhou (Associate Professor of Transportation Planning and Management at Shanghai Jiao Tong University), Dr. Wenxiang Li (Associate Professor of Transportation Planning and Management at the University of Shanghai for Science and Technology), and Dr. Jixiang Liu (Assistant Professor of Urban and Rural Planning at Xiamen University). The keywords for this Special Issue include transit-oriented development (TOD), transport policy, land-use policy, accessibility, mobility, travel behavior, shared mobility, built environment, and physical environment.

This Special Issue focuses on the interaction between transport and land use for sustainable urban planning and development. It offers a platform to share the latest accomplishments and research findings concerning the interaction between transport and land use, especially those with profound theoretical, methodological, and practical implications and those focusing on new theories, data, variables, models, sites, and their associated implications in the realm of transport and land-use interactions. This Special Issue is designed to be a catalyst for informed discourse and innovative advancements in the field. By addressing the multifaceted aspects of the relationship between transport and land use, we seek to contribute a comprehensive body of knowledge that can inform and guide sustainable urban planning practices.

After approximately thirty submissions underwent the standard review process, thirteen papers were selected for this Special Issue. The acceptance rate was around 40%. Moreover, the inclusivity of this Special Issue is reflected in the international scope of its authors. They are from ten countries, namely, China, the United States, the United Kingdom, The Netherlands, Germany, the Republic of Korea, Pakistan, Thailand, South Africa, and Saudi Arabia. This global representation not only underscores the widespread interest in the intersection of transport and land use for sustainable urban planning and development but also highlights the collaborative and cross-cultural nature of contemporary research endeavors.

The collected papers cover a wide range of research topics. Traditional topics include the impacts of transport on land use (Contribution 1), the impacts of land use on transport (more broadly, activity) (Contributions 2 and 3), transit-oriented development (TOD) (Contribution 4), and transportation planning and management (Contribution 5). New topics include, but are not limited to, autonomous vehicles (Contribution 6), evaluating the potential of CO_2 emission reductions for carbon neutrality (Contribution 7), and roadmaps for future mobility development (Contribution 8). Moreover, the techniques utilized include regression analysis (Contributions 1-3, 9, 10), principal component analysis (Contribution 11), mathematical programming (Contribution 12), simulations (Contribution 13), foresight and/or scenario analysis (Contributions 7 and 8), Bayesian networks (Contribution 5), and qualitative approaches (e.g., semi-structured interviews) (Contributions 4 and 6). The diversity in the collected papers closely mirrors the high level of attention paid to sustainable transport and land use by researchers worldwide in many fields, such as urban and rural planning, transportation engineering, and transport geography. Furthermore, all thirteen papers included in this collection fall under the category of "Original Research". Regrettably, this Special Issue includes no literature reviews.

3. Future Research Directions

The future research directions for sustainable transport and land use cover a broad spectrum, including, but not limited to, the implications of emerging mobility (mainly comprising shared mobility, autonomous vehicles, and electric vehicles), the impacts of emerging mobility on TOD, smart growth and the integration of transport and land use, transit development and urban/community renewal, intercity commutes via high-speed railways and regional integration, accessibility modeling and planning, climate-resilient transport infrastructure and land-use solutions, the interplay between transport emissions and land use, the application of cutting-edge data and analytical technologies, and operational integrated land-use transport models [1].

Among the future research directions suggested, three areas are particularly worthy of investigation. The first pertains to the implications of emerging mobility [5]. It is widely believed that emerging mobility could potentially help improve transport efficiency and sustainability by reducing the vehicle miles traveled, energy use, required vehicle fleet size, service costs, and emissions, and could limit the emissions of greenhouse gases and non-CO_2 pollutants in the transport sector [6–8]. That is to say, emerging mobility could offer significant benefits if implemented and managed properly. However, relying solely on the theoretical advantages of emerging mobility without considering the real-world effects could be a misstep. For example, while emerging mobility has great potential for reducing trip costs through energy savings, cheaper trips may induce trips and spur excessive vehicle use. These types of counterbalances or trade-offs indicate that emerging mobility may not always have positive environmental impacts. All in all, a deeper understanding of the implications of emerging mobility is beneficial in developing management strategies and supporting infrastructure (including land use) to help achieve the benefits of emerging mobility while mitigating any negative implications.

The second aspect of future research involves addressing climate change through the effective planning and management of transport and land use (i.e., seeking climate-resilient transport infrastructure and land-use solutions). Climate change profoundly influences ecosystems, weather patterns, sea levels, and human societies. It is the center of substantial attention from numerous stakeholders in today's society, evidenced by various national and local goals, such as carbon emission peak by 2030 and carbon neutrality by 2060 in China, carbon neutrality by 2045 in Sweden, carbon neutrality by 2045 in New Zealand, and carbon neutrality by 2025 in Copenhagen, Denmark. Addressing climate change through the effective planning and management of transport and land use is crucial, given that they are major contributors to greenhouse gas emissions. For example, the transport sector alone contributes to a quarter of total emissions [9]. Hence, sustainable practices in transport and land-use planning and management can play a significant role in mitigating climate change.

The third key research direction involves the application of cutting-edge data analytics to transport and land-use research. These data analytics have already revolutionized the landscape of transport and land-use research, offering unprecedented insights and transformative solutions. In recent years, advancements in collection methods for big/new data (e.g., global positioning systems, mobile phone signaling, and social media) have allowed urban/transport researchers and planners to gather detailed (even real-time) information about transport patterns, land use, and their dynamics. Moreover, cutting-edge data analytics, such as machine learning (more broadly, artificial intelligence), IoT (Internet of Things) analytics, and natural language processing, have emerged as powerful tools for processing large and complex datasets and extracting meaningful insights and patterns. These technologies enable researchers to analyze complex interactions between transport and land use, thereby facilitating a deeper understanding. Furthermore, the application of cutting-edge data analytics to transport and land-use research should not be regarded as merely "putting old wine into new bottles", as the Chinese proverb says. Instead, these analytics significantly enhance our capacity to comprehend the complexities of urban systems and contribute to evidence-based decision making. As cities continue to evolve and face the challenges of rapid urbanization, climate change, and resource constraints, applying cutting-edge data analytical techniques in research becomes paramount for creating innovative solutions that foster sustainable, inclusive, and resilient urban development.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: I thank the authors' creative concepts and dedicated endeavors, along with the valuable input provided by the other four Guest Editors (Yuanyuan Guo, Yaoming Zhou, Wenxiang Li, and Jixiang Liu) and the reviewers.

Conflicts of Interest: The author declares no conflict of interest.

List of Contributions

1. Chen, K.; Lin, H.; Liao, L.; Lu, Y.; Chen, Y.-J.; Lin, Z.; Teng, L.; Weng, A.; Fu, T. Nonlinear rail accessibility and road spatial pattern effects on house prices. *Sustainability* **2022**, *14*, 4700.
2. Liu, C.; Zuo, X.; Gu, X.; Shao, M.; Chen, C. Activity duration under the COVID-19 pandemic: A comparative analysis among different urbanized areas using a hazard-based duration model. *Sustainability* **2023**, *15*, 9537.

3. Im, J.-H.; Jang, S. Self-sufficiency of new administrative capitals (NACs) based on types and commuting characteristics of citizens: Case study of Sejong. *Sustainability* **2022**, *14*, 13193.
4. Almatar, K.M. Transit-oriented development in Saudi Arabia: Riyadh as a case study. *Sustainability* **2022**, *14*, 16129.
5. Xue, J.; Wang, J.; Yi, J.; Wei, Y.; Huang, K.; Ge, D.; Sun, R. Optimal parking path planning and parking space selection based on the Entropy Power Method and Bayesian Network: A case study in an indoor parking lot. *Sustainability* **2023**, *15*, 8450.
6. Chen, S.; Zong, S.; Chen, T.; Huang, Z.; Chen, Y.; Labi, S. A taxonomy for autonomous vehicles considering ambient road infrastructure. *Sustainability* **2023**, *15*, 11258.
7. Dong, J.; Li, Y.; Li, W.; Liu, S. CO_2 Emission reduction potential of road transport to achieve carbon neutrality in China. *Sustainability* **2022**, *14*, 5454.
8. Gerdsri, N.; Sivara, K.; Chatunawarat, C.; Jaroonjitsathian, S.; Tundulyasaree, K. Roadmap for future mobility development supporting Bangkok urban living in 2030. *Sustainability* **2022**, *14*, 9296.
9. Masoumi, H.; Aslam, A.B.; Rana, I.A.; Ahmad, M.; Naeem, N. Relationship of residential location choice with commute travels and socioeconomics in the small towns of South Asia: the case of Hafizabad, Pakistan. *Sustainability* **2022**, *14*, 3163.
10. Wang, S.; Lu, F.; Wei, G. Direct and spillover effects of urban land expansion on habitat quality in Chengdu-Chongqing Urban Agglomeration. *Sustainability* **2022**, *14*, 14931.
11. Lian, X.; Li, D.; Di, W.; Oubibi, M.; Zhang, X.; Zhang, S.; Xu, C.; Lu, H. Research on influential factors of satisfaction for residents in unit communities—Taking Ningbo City as an example. *Sustainability* **2022**, *14*, 6687.
12. Yao, J.; Xie, B.; Wu, X.; Zhang, C. Two-level programming model based on cooperative operation study of stakeholders in hazardous chemical storage. *Sustainability* **2023**, *15*, 1221.
13. Yao, J.; He, J.; Bao, Y.; Li, J.; Han, Y. Study on freeway congestion propagation in foggy environment based on CA-SIR model. *Sustainability* **2022**, *14*, 16246.

References

1. Wegener, M. Overview of land use transport models. In *Handbook of Transport Geography and Spatial Systems*; Emerald Group Publishing Limited: Bingley, UK, 2004; pp. 127–146.
2. Schiller, P.L.; Kenworthy, J. *An Introduction to Sustainable Transportation: Policy, Planning and Implementation*; Routledge: London, UK, 2017.
3. Litman, T.; Steele, R. *Land Use Impacts on Transport*; Victoria Transport Policy Institute Canada: Victoria, BC, Canada, 2017.
4. Yang, L.; Ao, Y.; Ke, J.; Lu, Y.; Liang, Y. To walk or not to walk? Examining non-linear effects of streetscape greenery on walking propensity of older adults. *J. Transp. Geogr.* **2021**, *94*, 103099. [CrossRef]
5. Liu, C.J.; Chen, K.; Bao, Z.; Ng, S.T.; Zhang, C.; Jiang, Z.-H. Assessing the impacts of connected-and-autonomous vehicle management strategy on the environmental sustainability of urban expressway system. *Sustain. Cities Soc.* **2023**, *99*, 104904. [CrossRef]
6. Santi, P.; Resta, G.; Szell, M.; Sobolevsky, S.; Strogatz, S.H.; Ratti, C. Quantifying the benefits of vehicle pooling with shareability networks. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13290–13294. [CrossRef]
7. Vazifeh, M.M.; Santi, P.; Resta, G.; Strogatz, S.H.; Ratti, C. Addressing the minimum fleet problem in on-demand urban mobility. *Nature* **2018**, *557*, 534–538. [CrossRef]
8. Ou, Y.; Bao, Z.; Ng, S.T.; Song, W. Estimating the effect of air quality on bike-sharing usage in Shanghai, China: An instrumental variable approach. *Travel Behav. Soc.* **2023**, *33*, 100626. [CrossRef]
9. United Nations. Transport Transformation Critical to Address Climate Change and Universal Access to Safe, Affordable, Resilient Mobility. 2021. Available online: https://www.un.org/sustainabledevelopment/blog/2021/10/transport-transformation-critical-to-address-climate-change-and-universal-access-to-safe-affordable-resilient-mobility/ (accessed on 3 November 2023).

sustainability

Article

A Taxonomy for Autonomous Vehicles Considering Ambient Road Infrastructure

Sikai Chen [1,*], Shuya Zong [2], Tiantian Chen [3], Zilin Huang [1], Yanshen Chen [4] and Samuel Labi [5]

1 Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA; zilin.huang@wisc.edu
2 Microsoft Software Technology Center Asia, Suzhou 215123, China; shuyazong@microsoft.com
3 Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Yuseong-gu, Daejeon 34051, Republic of Korea; nicole.chen@kaist.ac.kr
4 China Academy of Urban Planning and Design, Beijing 100044, China; chenys1999@sina.com
5 Center for Connected and Automated Transportation, Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA; labi@purdue.edu
* Correspondence: sikai.chen@wisc.edu

Abstract: To standardize definitions and guide the design, regulation, and policy related to automated transportation, the Society of Automotive Engineers (SAE) has established a taxonomy consisting of six levels of vehicle automation. The SAE taxonomy defines each level based on the capabilities of the automated system. It does not fully consider the infrastructure support required for each level. This can be considered a critical gap in the practice because the existing taxonomy does not account for the fact that the operational design domain (ODD) of any system must describe the specific conditions, including infrastructure, under which the system can function. In this paper, we argue that the ambient road infrastructure plays a critical role in characterizing the capabilities of autonomous vehicles (AVs) including mapping, perception, and motion planning, and therefore, the current taxonomy needs enhancement. To throw more light and stimulate discussion on this issue, this paper reviews, analyzes, and proposes a supplement to the existing SAE levels of automation from a road infrastructure perspective, considering the infrastructure support required for automated driving at each level of automation. Specifically, we focus on Level 4 because it is expected to be the most likely level of automation that will be deployed soon. Through an analysis of driving scenarios and state-of-the-art infrastructure technologies, we propose five sub-levels for Level 4 automated driving systems: Level 4-A (Dedicated Guideway Level), Level 4-B (Expressway Level), Level 4-C (Well-Structured Road Level), Level 4-D (Limited-Structured road Level), and Level 4-E (Disorganized Area Level). These sublevels reflect a progression from highly structured environments with robust infrastructure support to less structured environments with limited or no infrastructure support. The proposed supplement to the SAE taxonomy is expected to benefit both potential AV consumers and manufacturers through defining clear expectations of AV performance in different environments and infrastructure settings. In addition, transportation agencies may gain insights from this research towards their planning regarding future infrastructure improvements needed to support the emerging era of driving automation.

Keywords: autonomous vehicles; automated driving; society of automotive engineers; road infrastructure; operational design domain; taxonomy

Citation: Chen, S.; Zong, S.; Chen, T.; Huang, Z.; Chen, Y.; Labi, S. A Taxonomy for Autonomous Vehicles Considering Ambient Road Infrastructure. *Sustainability* **2023**, *15*, 11258. https://doi.org/10.3390/su151411258

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 20 June 2023
Revised: 13 July 2023
Accepted: 14 July 2023
Published: 19 July 2023

1. Introduction

Autonomous vehicles (AVs) have garnered substantial attention from various sectors, including government agencies and policymakers, the automotive and technology industries, and academia, due to their potential to enhance road safety, improve travel efficiency, and reduce energy consumption [1–4]. In several countries, such as China, Japan, the United Kingdom, Sweden, Germany, and Japan, AVs have been permitted or are in the process of being permitted for deployment in specific locations such as experimental test tracks, academic campuses, and demarcated urban zones, with restrictions to avoid mass public exposure [5–7]. In the U.S., in 2021, President Biden signed the USD 1 trillion Infrastructure Investment and Jobs Act (IIJA), which commits $110 billion to roads, bridges, and other significant projects, and mandated the United States Department of Transportation (USDOT) to develop the Automated Vehicles Comprehensive Plan (AVCP) to ensure American leadership in autonomous vehicle technologies [8]. The Federal Highway Administration (FHWA) has developed the innovative CARMA Platform to foster collaboration aimed at enhancing transportation efficiency and safety [9]. In addition, the American Association of State Highway and Transportation Officials (AASHTO) has published a document titled "Connecting on CAVs", featuring newly established policy principles for connected and autonomous vehicles. These principles are intended to facilitate the advancement of connected and autonomous vehicle (CAV) technology development [10].

Despite the growing interest in AVs, there seems to exist a lack of consensus on the definition of the term AV. This conundrum relates to the purpose of this paper and motivates a closer study of the reasons for such lack of taxonomical consensus. Automation generally refers to the use of control systems for operating equipment or performing human tasks. Therefore, vehicle automation is part of a broader trend toward the replacement of humans in various functions, including machine operation. This trend is driven by reasons that include (i) the rapid growth of information and communication technologies (ICT), which has led to increased computing power, and (ii) the need for greater efficiency and improved operator safety. Cognitive psychologist Lisanne Bainbridge pointed out that, paradoxically, with automation, human involvement becomes more critical, even as humans become less involved in the operation of automated systems [11]. Automation involves not only control systems, computer engineering, and ICT, but also psychology, social sciences, and business. As such, it is not surprising that pilot research efforts on the deployment efficacy of automated driving systems and their policy development are quite multidisciplinary in nature as they involve knowledge from these diverse fields. It appears that the multidisciplinary nature of automation has been a curse as much as it has been a blessing, because to date, there seems to be a longstanding lack of consensus on the definition of the term "autonomous vehicles" [12].

The Society of Automotive Engineers (SAE) has developed a taxonomy scale to standardize the levels of driving automation and related terminology for the benefit of the automotive industry and AV policymakers. The scale consists of six levels of vehicle automation, ranging from Level 0 to Level 5 [13]. Level 0 vehicles are fully controlled by the driver, while Level 1 vehicles have a single automated system for assisting with either steering or cruise control. Level 2 vehicles, such as Tesla Autopilot and Cadillac Super Cruise, feature automation for both steering and cruise control. In general, drivers still perform a significant portion of the driving tasks in Level 0 to Level 2 vehicles. However, from Level 3 to Level 5, the automated driving systems take over all driving tasks when engaged. Level 3 vehicles can make informed navigation decisions and undertake driving tasks, but still require human intervention in certain situations. Level 4 vehicles can intervene, if necessary, but are limited to a specific geographic area. Finally, Level 5 vehicles represent the pinnacle of automation and do not require any human intervention under any circumstances.

The SAE taxonomy provides a clear and comprehensive description of the driving automation levels based on the capabilities of the automated system. However, it lacks information on the necessary infrastructure to support each level of automation [14,15]. In reality, the road network is composed of diverse road types, requiring AVs to seamlessly transition between them during their operation. Our main insight is that for each 4-x level, different types of infrastructure, i.e., capabilities, are required. In this paper, we first define the different types of infrastructure, which are classified based mainly on the characteristics of the infrastructure/environment:

(1) Dedicated Guideways:
- Lanes are exclusive and fully controlled;
- Intelligent and complete infrastructure is accessible;
- Other road users such as pedestrians seldom occur.

(2) Expressways:
- Reliable V2I and V2V communication is accessible;
- Surrounding vehicles are moving in the same direction;
- Other road users seldom occur. Wild animals may suddenly appear but at a relatively low frequency.

(3) Well-Structured Roads:
- Clear lane markers and complete traffic signals are accessible;
- Smart and communicable infrastructure may be inaccessible;
- A large number of other road users such as pedestrians and bicycles exist.

(4) Limited-Structured Roads:
- Road lane markers and traffic signs are incomplete or even unavailable;
- Intelligent infrastructure is usually inaccessible;
- The road may be covered by flood, ice, or dirt such that lane markers are invisible;
- Some wild animals, pedestrians, vehicles, and other road users exist in the surroundings.

(5) Disorganized Areas:
- The surroundings are constituted by huge crowds of people, bicycles, motors, and other road users;
- Space suitable for driving is usually limited;
- Assistance from nearby intelligent infrastructure is inaccessible.

Figure 1 depicts a bird's-eye view of a hypothetical road network that comprises a variety of road types: dedicated guideways for AVs (labeled with '①' at the top right), expressways (labeled with '②' at the top left), well-structured roads (labeled with '③' in the middle), rural roads as limited-structured roads (labeled with '④' at the bottom left), and disorganized areas in communities (labeled with '⑤' at the bottom right). The tunnel in '①' is an example of a dedicated guideways for AVs, with advanced infrastructure that is capable of guiding AVs and protecting them from unexpected interactions with pedestrians, bicyclists, animals, and so on. When driving in this tunnel, the vehicle does not need navigation capabilities or real-time object detection capabilities. Since the vehicle does not require human take-over, vehicles without any capabilities of navigation capabilities or real-time object detection but have simple self-driving functions (e.g., lane-keeping) to drive in such tunnels can be identified as L4 in the SAE classification framework. However, such vehicles (without any capabilities of navigation capabilities or real-time object detection) will not be able to realize L4 automation on the roads labeled with '③', which are well-structured urban roads with mixed traffic. An SAE L4 vehicle must have the capabilities of real-time object detection and collision avoidance when driving in such an area (due to complex interactions with human-driven vehicles, pedestrians, etc.). Obviously, there is a clear distinction between these two types of vehicles to realize L4 automation in areas '①' and '③'.

Now, consider the expressways labeled as '②'; these roadways can have higher speed limits and fewer unexpected obstacles compared to urban areas. However, they may still encounter unexpected situations, like sudden traffic congestion or a vehicle breakdown. Thus, an SAE L4 vehicle on an expressway still needs to be capable of real-time object detection, collision avoidance, and the ability to navigate through the potentially high-speed traffic. While both the '②' expressways and '③' well-structured urban roads require an SAE L4 vehicle to have the capabilities of real-time object detection and collision avoidance, the primary difference lies in the traffic dynamics. Thus, an L4 AV must be capable of handling high-speed decision-making on expressways, and at the same time, it must deal with the complexity and diversity of urban traffic scenarios. Roads labeled as '④' represent rural roads as an example of limited-structured roads, where there may not be clear road markings or signals, and the vehicle may encounter unexpected obstacles like wildlife or agricultural vehicles. In such scenarios, an SAE L4 vehicle needs robust navigation capabilities and advanced perception to handle less predictable road conditions. Further, consider an AV classified as L4 (represented in red in the figure): while the red vehicle exits the tunnel and navigates urban roadways, particularly in the disorganized area labeled as '⑤', it must be aware of the changes in surrounding infrastructure to take appropriate actions accordingly; otherwise, catastrophic traffic accidents may occur.

Figure 1. A bird's-eye view of a simulated road network that comprises a variety of road types: dedicated guideways for AVs (labeled with '①' at the top right), expressways (labeled with '②' at the top left), well-structured roads (labeled with '③' in the middle), limited-structured roads (labeled with '④' at the bottom left), and disorganized areas (labeled with '⑤' at the bottom right).

Clearly, the full potential of automated driving systems can only be realized when accompanied by infrastructure that possesses the appropriate level of technological advancement or "smartness". Hence, it is vital to understand the crucial role that infrastructure plays in autonomous driving [16,17]. Without appropriate infrastructure support, an AV that is considered to be at a certain SAE level in certain areas may not realize a similar level of autonomous driving performance in other areas.

2. Study Objectives and Organization of the Paper

It is expected that the deployment of AVs will follow a gradual progression, commencing with simple operational domains such as dedicated lanes, access-controlled freeways, and rural arterials, before moving on to more complex domains including urban arterials, intersections, and city/town streets. Both prior to the onset of this progression and at each subsequent stage, there is a need to comprehend the crucial role that infrastructure plays in supporting automated driving systems. This can be achieved through incorporating the requisite level of environmental infrastructure support required for each level of automation in the established SAE taxonomy.

However, as previously discussed in this paper, the current definition of SAE Level 4 lacks distinction between the varying capabilities of vehicles with this level of automation across different environment–infrastructure domains. To throw more light and stimulate discussion on this issue, this study proposes an enhancement of the current taxonomy through incorporating the role played by infrastructure in automated driving systems, specifically with regard to Highly Automated Driving Systems (HADS)—Level 4. The focus on Level 4 systems is due to their greater potential for near-future implementation compared to Level 5, as well as the recognition that HADS can only attain full operational capability under specific and limited conditions that require a clear understanding of the necessary infrastructure.

The major contributions of this work are as follows:

- With our proposed supplemental taxonomy, various driving conditions can be classified, enabling the vehicle to fully understand its ambient driving environment. This extends beyond alerting the vehicle when the level of infrastructure advancement changes from Level 4-A to Level 4-B. Indeed, it underscores the essential need for an AV to be adaptable across all Level 4-x infrastructures. In this context, our proposed taxonomy does not merely serve as a warning system but also functions as a foundation upon which AVs can evaluate and adjust their capabilities accordingly. This adaptability is crucial in ensuring that human intervention is only sought when necessary, thereby maximizing the autonomy of these vehicles. Furthermore, the proposed taxonomy will also create an industry standard that helps manufacturers and vehicle sellers clearly demonstrate the capabilities of their AVs. Instead of simply advertising a "Level 4" vehicle, manufacturers should explicitly inform consumers about the specific levels of road infrastructure environments where the Level 4 vehicle is capable of operating safely and effectively.
- The proposed supplement to the SAE taxonomy incorporates the role of the environment–infrastructure domain, offering a more comprehensive approach to characterizing Level 4 automated driving systems. This supplement aims to provide clarity regarding potential subsets of Level 4 automation, enhance the implementation of the SAE taxonomy, and serve as a reference for the design, testing, and evaluation of high-level AVs for the benefit of all stakeholders. Through directly characterizing the operations of HADS in terms of both automation level and infrastructure smartness, this supplement has the potential to create realistic expectations, increase confidence, and enhance the credibility of autonomous vehicle operations.
- This study has the potential to provide a robust foundation upon which AV manufacturers can precisely delineate the operational capacities of their vehicles. Moreover, it can enable AV users to attain a more accurate understanding of their vehicle's capabilities. Furthermore, the proposed supplement can inform government regulators and policymakers in formulating suitable policies and regulations that consider the infrastructure–environment domain. Additionally, it can provide road agencies with the necessary information to make informed decisions regarding investments in infrastructure to support AV operations. Ultimately, the proposed supplement can offer infrastructure managers, investors, and policymakers stronger justifications for policies, initiatives, and investments aimed at preparing infrastructure for AVs.

The organization of the rest of this paper is as follows: Section 3 provides a brief overview and critique of the existing taxonomy system for AVs. In Section 4, we examine the various factors that impact AV operations. The supplementary taxonomy is presented in Section 5, followed by a consideration of the future development of AVs in Section 6.

3. Review of the Current SAE Taxonomy and Its Limitations

3.1. Existing SAE Taxonomy

Several institutions have established automated vehicle classifications, including the International Organization of Motor Vehicle Manufacturers (OICA) [18], the Germany Federal Highway Research Institute [19], and the SAE [13]. Of these, the SAE's taxonomy, consisting of six levels of vehicle automation, has garnered the most widespread acceptance and has been adopted as the industry standard. This system has been endorsed by the US Department of Transportation [13,20]. The current SAE taxonomy, although widely adopted and recognized as the industry standard, does have certain limitations. One such limitation is that the taxonomy primarily focuses on the technical aspects of vehicle automation, such as the level of driver involvement and the extent of automation, without considering the variability in driving conditions and environments across different regions, which can have a significant impact on the performance and reliability of automated systems [21–26].

To effectively characterize and evaluate the complex nature of automated driving systems and related issues, various stakeholders, including the SAE and the USDOT, have played a crucial role in shaping key terms and concepts. One such concept is the operational design domain (ODD), which serves to describe the specific conditions and scenarios under which an automated system or feature is designed to operate. The ODD incorporates environmental, geographical, time-of-day, and roadway restrictions and characteristics, as outlined in the literature [4,13]. It is worth noting that the elements of the ODD can be classified into two broad categories: those that are within the direct control of the relevant agency, such as infrastructure quality, and those that are beyond the agency's control, such as weather conditions (e.g., ice, wind, fog, smoke) and traffic conditions (e.g., occlusion by large vehicles). The dynamic driving task (DDT), as defined by the SAE [13], encompasses all the functions required to operate a vehicle on-road in traffic. As detailed in Table 1, each level of automation within the SAE taxonomy has a distinct set of requirements that a vehicle must meet before it can be considered operational at that level.

Table 1. SAE Taxonomy [13].

Level	**Name**	**Narrative Definition**	**DDT**		**DDT Fallback**	**ODD**
			SLLVMC	**OEDR ***		
Driver Performs Part or All of the DDT						
0	No Driving Automation	The performance by the driver of the entire DDT, even when enhanced by active safety systems.	Driver	Driver	Driver	n/a
1	Driver Assistance	The sustained and ODD-specific execution by a driving automation system of either the lateral or the longitudinal vehicle motion control subtask of the DOT (but not both simultaneously with the expectation that the driver performs the remainder of the DDT.	Driver and System	Driver	Driver	Limited
2	Partial Driving Automation	The sustained and ODD-specific execution by a driving automation system of both the lateral and longitudinal vehicle motion control subtasks of the DDT with the expectation that the driver completes the OEDR subtask and supervises the driving automation system.	System	Driver	Driver	Limited
ADS ("System") performs the entire DDT (while engaged)						
3	Conditional Driving Automation	The sustained and ODD-specific performance by an ADS of the entire DDT with the expectation that the DDT fallback-ready user is receptive to ADS-issued requests to intervene, as well as to DDT performance-relevant system failures in other vehicle systems and will respond appropriately.	System	System	Fallback-ready user (becomes the diver during fallback)	Limited

Table 1. *Cont.*

Level	Name	Narrative Definition	DDT		DDT Fallback	ODD
			SLLVMC	OEDR *		
	Driver Performs Part or All of the DDT					
4	High Driving Automation	The sustained ODD-special performance by an ADS of the entire DDT and DDT fallback without any expectation that a user will respond to a request to intervene.	System	System	System	Limited
5	Full Driving Automation	The sustained and unconditional (i.e., not ODD-special) performance by an ADS or the entire DOT and DDT fallback without any expectation that a user will respond to a request to intervene.	System	System	System	Unlimited

Note: SLLVMC—sustained lateral and longitudinal vehicle motion control. '*' denotes the OEDR—object and event detection and response. n/a—not applicable.

3.2. Role of Road Infrastructure

The role of infrastructure in supporting automated driving is significant [27–29]. It is expected that road infrastructure in the AV era will be equipped with smart features to augment the AV's capabilities. First, the infrastructure should enhance the sensing abilities of AVs through collecting, analyzing, and transmitting data regarding the driving environment, including the characteristics of roadways and other vehicles in the traffic stream. This information can be used to inform the strategic, tactical, and operational driving decisions of the AV [30]. Secondly, the infrastructure should provide support to AVs in navigating through challenging driving conditions, such as periods of impaired sensing or navigation capabilities caused by adverse weather conditions, occlusion by larger vehicles, or lack of connectivity [31]. The USDOT, in its Automated Vehicle Comprehensive Plan document, duly recognizes these prospective roles of AV infrastructure and the environment, and the document mentions that modernizing the regulatory environment and preparation of the existing transportation infrastructure is critical for successful automated transportation [8,32]. Also, the American Association of State Highway and Transportation Officials indicated that keeping a safe driving environment for AVs through defining regulations and constructing smart infrastructure is a major issue [33]. Notably, McAslan et al. [34] found that many regional planning agencies in the United States have implemented policies aimed at enhancing infrastructure maintenance to facilitate the testing and deployment of AVs. Thus, the importance of smart infrastructure in ensuring the safe and efficient operation of AVs cannot be overstated.

The significance of infrastructure support for the successful operation of AVs is highlighted by the fact that several AV manufacturers and technology companies invest in the development of dedicated infrastructure and accompanying hardware. One such example is the tunnel system developed by Elon Musk's The Boring Company in Los Angeles, known as the "Test Tunnel". This 1.2-mile-long tunnel, shown in Figure 2a, was designed for the purpose of research and development for Musk's innovative vision of a network of underground highways [35]. Within this controlled environment, autonomous vehicles have the potential to reach speeds of up to 25 mph. The "Future Bus" project developed by Mercedes-Benz involves the operation of buses on dedicated bus-only lanes that are equipped with vehicle-to-infrastructure (V2I) connectivity facilities, providing information on the bus route and station locations. This allows the bus to operate at a speed of 43 mph and make precise stops at designated stations, as shown in Figure 2b. Another example is the 3.9 km guideway used by the Heathrow Airport Authority in London for elevated automated personal transportation pods that utilizes a sled guideway, as shown in Figure 2c [36]. Even where testing is carried out on surface roads, the urban–rural nature of the highway is of paramount importance in designing road infrastructure to support AV operations. For example, for Volvo's autonomous truck, which runs on rural highways, there is access control infrastructure, well-marked pavements, and a relatively uniform traffic environment. In contrast, Google's self-driving cars were tested in urban areas that had no access control and had more dynamic and challenging driving scenarios including several traffic signals, dense traffic, and traffic jams. As we can see, the role of infrastructure is not only to provide a physical platform for AVs, but also to enable the achievement of autonomous driving in specific areas through reducing the need for complex real-time environmental perception and decision-making capabilities.

Another noteworthy example is the Ultimate Urban Circulator Program (U2C) in Jacksonville, Florida [37]. Through repurposing the existing downtown circulator guideway for autonomous vehicles and combining this with flexible routes outside the downtown area, Jacksonville showcases a real-world example of how cities can effectively adapt existing infrastructure to usher in an era of autonomous transportation.

(**a**) (**b**) (**c**)

Figure 2. Infrastructure and supporting hardware for AVs invested by technology companies. (**a**) The Test Tunnel [35]; (**b**) Mercedes-Benz Bus of the Future [38]; (**c**) Heathrow ULTra pods [39].

It is interesting to note that both Volvo's autonomous trucks and Google's self-driving cars both fall into Level 4 according to SAE's definition. However, their infrastructure requirements for their successful operation are very dissimilar: Volvo's autonomous trucks cannot operate in the driving environment of Google's self-driving cars, and vice versa. The current SAE taxonomy for Level 4 does not address the distinction between these infrastructures and environments. In describing an automated driving system, it seems reasonable to mention the functionality and requirements of off-board infrastructure conditions and the corresponding driving environment. Scholars argue that a fully automated driving system is attainable only in the context of their constraints and that the current SAE taxonomy system for vehicle automation should be further described with specificity to a prevailing condition, instead of broad categorization that is irrespective of infrastructure conditions [23]. Ran et al. [40] proposed a definition of intelligence in smart road infrastructure called connected automated highway (CAH). They defined CAH using five levels and illustrated how the levels function with the CAV in collaborative automated driving system (CADS). Tengilimoglu et al. [26] interviewed 168 experts from 29 countries to determine the infrastructure-related requirements for safe operation of SAE Level 4 autonomous vehicles regarding (1) deployment paths, (2) autonomous driving road certification, (3) key infrastructure elements, and (4) factors impacting safe operation. Saeed et al. [41] developed a road infrastructure classification and discussed the challenges and opportunities related to the readiness of infrastructure for connected and autonomous vehicles (CAVs). Soteropoulos et al. [42] proposed a framework and relevant metrics for evaluating autonomous driving performance based on the relationship between the current technological state of autonomous driving systems (ADSs) and various domains.

Moreover, it is crucial for prospective users of AVs of a certain level of automation, to be aware of the level of infrastructure advancement that is consistent with their AV. If this is not done, the high-level AV user may be operating their vehicle in a low-level environment/infrastructure, thereby jeopardizing the travel efficiency and safety of the AV user and other road users. For example, a purchaser of a "Level 4" autonomous vehicle should be made aware that the vehicle can attain its full potential only under specific strictly designed infrastructure which matches that level of automation and not in any normal road environment–infrastructure domain. The questions that arise, then, are as follows: realizing that the capabilities of highly automated driving systems are different under different levels of infrastructure advancement, does the specification of the autonomous vehicle's level of automation alone suffice in characterizing its operational capabilities? Should a complete characterization of operational capability not depend on both the level of vehicle automation as well as the level of the environment–infrastructure domain?

3.3. Limitations of SAE Taxonomy

The SAE taxonomy system continues to be the most prevalent and cited reference in the field of automated driving. However, the authors of this paper believe are a few opportunities to improve the SAE taxonomy.

3.3.1. From the ODD Perspective

The SAE taxonomy does not incorporate any ODD elements to accompany the levels of automation, particularly the higher levels. The extensive variety of the infrastructure and environmental conditions (including the level of connectivity, the level of intelligence of supporting infrastructure, and the quality levels of AV-supporting road facilities including lane markings) are not addressed in the SAE's taxonomy. In this regard, it can be argued that the level of infrastructure advancement or "smartness" needed to support any given level of automation is a key consideration in assessing the capability of an autonomous vehicle [43].

As discussed in a previous Section 3.2 of this paper, facilities such as the Test Tunnel [35], Mercedes-Benz Future Bus, and Heathrow ULTra pods [36] all enable Level 4 automated driving but accomplished this only under very specific infrastructure conditions. In such carefully designed environments and infrastructure, Level 4 performs well. However, the same Level 4 vehicles are unlikely to exhibit such performance at other less-defined conditions of the environment and infrastructure, for example, a typical urban environment with errant HDVs, pedestrians, and other uncertainties in the traffic stream and the roadway environment in general. Therefore, the performance of Level 4 autonomous vehicles is expected to be different in different environments.

3.3.2. From the DDT Perspective

There is also an argument to be made for supplements to the SAE taxonomy, from the DDT perspective. The infrastructure required to support an automated driving system is not described when defining a high-level automated driving system. One fundamental assumption adopted by SAE is that DDT is performed by either in-vehicle systems or human drivers. Nevertheless, supported by the rapid advancement of wireless communication, intelligent infrastructure has played a significant role in DDT performance. For example, recent research studies have extensively deployed vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication (or V2X in short) to facilitate the monitoring of the roadway environment, trajectory planning, and decision-making in automated driving [14,15,44,45]. A practical illustration is the PATH project, where cars are driven automatically at close spacing on a freeway with magnetic markers embedded under the road. Through decreasing the spacing of cars, the density of traffic on a highway can be increased without requiring additional lanes [46]. In more recent studies on platoon-based cooperative driving, a vehicle collects kinetic data on neighboring vehicles through V2V communication and maintains its state in a cooperative way [47,48]. This is evident not only in the key role of infrastructure in achieving highway performance but also in the need to recognize the limitations of road infrastructure in automated driving and accounting for such limitations in characterizing the capabilities of automated driving systems.

Last but not least, policymakers and the public need more explicit information regarding scenarios in which a specific type of HADS can operate safely and regarding the infrastructure requirement associated with a given level of performance of the automated system [26,34,41]. Several companies have stated that they have developed vehicles with Level 4 automation [49,50]. However, as discussed previously, some projects are launched only in strictly controlled conditions, and significant variations in the capability of Level 4 automated driving systems discrepancies have been observed. This would not only lead to public confusion about the capability of their high-level autonomous vehicles but also increase the agency's cost of road systems management because software and hardware requirements vary widely for successful automated driving in different environments. In making new policies or passing new traffic legislation that duly recognizes the limitations and capabilities of automated driving systems in different environments and infrastructures, policymakers need clearer ways to communicate with AV companies. For these reasons, there seems to be some merit in the notion that it is useful to identify the requisite infrastructure conditions to support each level of automated driving, and that such a description could serve as a supplement to the existing SAE taxonomy.

4. Roadway Conditions and Facilities That Influence the Performance of Highly Automated Driving Systems

Given the highly dynamic nature of not only traffic operations and conditions but also the road environment, safe driving can be a demanding task. The operator needs to be able to interact effectively with other vehicles in its proximity and elements in the wider roadway environment including roadside units, traffic signs, pedestrians, two-wheelers, road surface conditions (ice, potholes), atmospheric conditions (fog, rain), and so on. For these reasons, driver licensing departments in most countries require that the driver should have adequate mental and visual acuity, and intuitiveness. In this emerging age of autonomous transportation where humans prepare to hand over the driving task to automation, one of the dilemmas that remain is how to train the automated systems to acquire the full capabilities of a "good" human driver [51–56]. Fortunately, researchers have identified a number of roadway conditions and facilities that influence the performance of highly automated driving systems. With this knowledge, efforts can be made to identify the extent to which an environment and infrastructure combination is deemed deficient for automated driving and to serve as a basis for investments to upgrade the environment and/or infrastructure. Also, such knowledge can help support the argument that automated systems classifications should address not only the vehicle capabilities but also the reinforcing or extenuating conditions of the environment and/or infrastructure in which the system is in operation. In Section 4, we delve into these various aspects under three subsections: Section 4.1, Traffic Conditions; Section 4.2, Cyber Infrastructure; and Section 4.3, The Language of the Road. Each subsection aligns with our proposed Level 4-x infrastructure sublevels and details the specific infrastructure prerequisites required for each. It is important to note that our intent here is to provide an integrated understanding of the complex interactions between automated vehicles, traffic conditions, cyber infrastructure, and road environments, and how these interactions relate back to our proposed Level 4-x taxonomy.

4.1. Traffic Conditions

The safety of automated driving is particularly challenging in complex and dynamic roadway traffic conditions [57]. The Waymo Safety Report underscores this perspective [58]. It discloses findings from one million miles of autonomous driving tests conducted on public roads in California and Arizona, emphasizing the profound influence of environmental and infrastructure factors on the overarching safety of autonomous driving systems. Roadway environments that are auspicious for automated driving are rare in certain areas where traffic flow is non-uniform, traffic composition is heterogenous, and the flow of multitype traffic is chaotic. For example, in several cities in developing countries, two-wheelers constitute over half of the vehicle distribution, resulting in unstable traffic flow [59,60].

Also, there is the issue of non-uniform traffic regulations across jurisdictional boundaries (for example, switching between left-hand vs. right-hand traffic) to which human drivers generally adapt quickly. In such situations, a highly automated driving system may encounter significant difficulty and, therefore, will require advanced sensing and tracking capabilities. Furthermore, urban roadway environments are typically characterized by non-verbal communication among the users of the roadway space. There exists a myriad of movement-related cues and gestures that automated systems may not comprehend [61]. For example, an AV will need to recognize a person standing at any part of the roadway waving their hands and ascertain whether that person is a traffic police official directing traffic, a passenger hailing a taxi, or a pedestrian requesting vehicles to stop so that other vulnerable road users may cross the street.

The implementation of automated driving in complex and intricate urban environments presents considerably more formidable challenges than doing so in well-defined and relatively conducive environments. Therefore, it is obvious that the automated system cannot reach its full potential in a non-conducive environment. To date, there is no agreed standard regarding the order of traffic conditions for automated driving. Moreover, significant variations in such conditions have been observed in pilot AV projects in recent decades, as shown in Table 2. The interested reader can find many more recent projects in [62]. Therefore, it can be argued that a classification system that specifies the conditions under which the AV will realize its full potential will be beneficial.

Table 2. Comparison of traffic conditions among AV projects (most up to date).

Project Name	Traffic Condition of Driving Scenario	Launch Year	Country	Latest Progress (2023)
Benz Future Bus	A dedicated bus-only lane of length is approximately 12 miles. Pedestrians seldom enter the lane.	2017	Germany	The Mercedes-Benz "Drive Pilot" system can only be used during daytime on the highway at speeds of up to 40 mph.
Volvo autonomous truck 'Vera'	A predefined and fixed route between logistic ports. Most of the roads are public roads with limited traffic.	2017	Sweden	The company's business remains focused on transporting products from logistics centers to ports
Robot Taxi	A 5 km road between Tokyo station and the Roppongi area. The on-road traffic is light and stable, but pedestrians are present at intersections and crossways.	2019	Japan	Autonomous vehicles intended for use as delivery robots or tour buses on routes in sparsely populated areas
Pony autonomous taxi	Crowded and busy roads in Guangzhou. Copious amounts of pedestrians and bicycle traffic at intersections.	2019	China	Received permission to run a fully automated driverless ride-sharing service in Guangzhou, China
Waymo self-driving	Waymo operates commercial self-driving taxi services in Phoenix, Arizona, and San Francisco, CA.	2020	USA	Waymo is now permitted to begin driverless taxi service in San Francisco, California, after receiving permission from the California Public Utilities Commission.

4.2. Cyber Infrastructure

Intelligent equipment installed on the infrastructure at or near the roadway (for example, on the roadway pavement, roadside structures, nearby buildings, drones, guardrails, and traffic signals) are typically capable of V2X communication and provide information that can influence the performance of the automated driving system. They have been used widely to facilitate localization, construct mapping, characterization of the roadway environment, and other functions for purposes of automated driving [63–66]. For example, a "bottleneck manager" typically installed at freeway locations with recurrent bottleneck congestion receives and prioritizes requests from AVs and optimizes their trajectories to reduce congestion and smooth traffic flow [53,67–71]. In addition, the concept of V2X communication has been extended to connected vehicle-to-pedestrian communication (V2P) via smartphone applications. V2P has been applied in several contexts of automated driving particularly, pedestrian safety enhancement through broadcasting locations and potential movements of pedestrians to the AVs in the disorganized area [72–74]. Also, Nikola Motor Company attempted to establish a connection between self-driving vehicles and pedestrians using the cell phones of individuals [75].

Highway agencies seek guidance on the changes in infrastructure design and management needed for AV operations because a robust network of physical and cyber infrastructure can help resolve several obstacles to automated driving. Ref. [41] discussed the challenges and opportunities associated with infrastructure preparation for AVs, identified stakeholder roles regarding AV infrastructure provision, and discussed uncertainties regarding AV market penetration and level of autonomy during the AV transition period. Concerning cyber infrastructure, there exist concerns about their practical deployment. First, the wide implementation of cyber infrastructure is not always feasible because these technologies hinge on the availability of wireless communication and big data computing capabilities, which are costly [76–78]. Second, effective communication hinges on the proper functioning of all nodes in the communication process. Malfunction of any component (for example, the communication device or a stable internet connection) could cause mischaracterization of the traffic environment, and in extreme cases, traffic accidents [61]. Therefore, HADS that depend partially on supporting infrastructure would be inadequate in areas where the quantity or quality of such infrastructure (software and hardware) is inadequate.

4.3. The Language of the Road

The "language of the road" is a term that collectively represents traffic lights, signals, signs, and road markings (road markings here exclude the physical characters painted on the road pavement surface). These infrastructures play an important role in traffic operations as they provide information to drivers, provide alarm of impending hazards or special attention downstream, and help in the navigation task [79,80]. AVs that can understand the language of the road will be able to enhance their traffic movements and trajectory planning [81]. From the literature, it is shown that attempts have been made to enable AVs to recognize road signs [82–84]; however, at the current time, the reliability of the developed recognition systems seems to be far from perfect. First, as the figures and text in Figure 3 show, the shapes and appearances of road language can be intricate and elaborate. Therefore, an AV may encounter difficulty distinguishing, for example, the road markings in the first two rows of Figure 3, which hold similar profiles but different meanings. Stop signs, which are typically encoded in the local language, have characters that are dissimilar across countries. Furthermore, the units of speed are not the same across countries: most countries, including China, Australia, and Singapore, use the metric system (i.e., kilometers per hour), while others such as the United States use the imperial system (i.e., miles per hour).

Secondly, in cases where the AV depends partially on mapped information established based on road inventory or where AV maneuvering algorithms are based on prior knowledge of infrastructure locations [85], any deviations from the mapped inventory and actual ground-truth inventory can threaten the AV's safety and travel efficiency. Mapping updates may be infrequent enough to capture some changes in the road inventory. For example, it can be challenging for automated systems facilities to comprehend or recognize the existence of temporary traffic control and management installed by the road agency or the police.

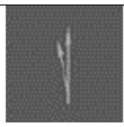
	Interpretation	Appearance	Colors	Shape	Country/Region
Road Markings	Lane opens to all vehicles at the end of the bus lane on the left		White	Arrows	Hong Kong
	Entrance to deceleration lane		White	Arrows	Hong Kong
Traffic Signs	Stop		Red, White	Octagon	China
			Red, Yellow	Octagon	Nigeria
			Red, White	Octagon	Chile
	No Parking		Red, Blue	Circle	Hong Kong
			Red, White, and Black	Circle	New Zealand
	Speed Limits		Black, White	Rectangular	United States (Unit: miles/h)
			Red, White, and Black	Circle	China (Unit: km/h)

Figure 3. Language of the road—examples.

Thirdly, the AV's ability to interact with supporting infrastructure may be jeopardized by a non-favorable locational or functional relationship with the infrastructure. For example, it may be difficult for AV sensors and image-detection algorithms to read traffic signs due to adversities related to the angle position of the signs, the sign retro-reflectivity [86–88], time of day (daytime glare or nighttime obscurity), inclement weather, and the sun's position (the AV's camera view may be positioned towards or away from the sun).

Finally, the typical urban skyline is strewn with several objects of various sizes, shapes, and colors, and it may be difficult to distinguish between these objects and those that were installed to support traffic operations, such as traffic lights and road signs. This places a large burden on sign-detection algorithms [89]. As such, understanding the language of the road can be generally more challenging in metropolitan areas compared with rural areas that have clear skylines.

4.4. Physical Characters "Of the Ground"

The physical characters often painted on the road pavement surface can influence the effectiveness of not only normal human driving [90] but also, automated driving. These include lane and road appearance and clarity, visibility, road curvature, and other characteristics. Most known AV demonstrations have been carried out on simple roads that offer a conducive environment for AV operations, for example, well-marked lanes and high standards of road design [89]. Recent research studies have adopted assumptions regarding the consistency of lane/road texture, lane/road width, alignment, and surface markings [91,92]. However, these conditions do not exist on relatively less structured roads where there is greater uncertainty in the design and operational conditions of roadways. Automated driving on unstructured roads is therefore challenging, and the specification of highly automated systems must duly account for the conditions under which HADS realize their full potential.

5. The Proposed Supplement for Classifying Highly Automated Driving Systems

As the preceding sections have established, the capability of automated driving systems is not the same in all environment/infrastructure domains, and therefore the HADS user must be aware of the capabilities of their system when it is operating in each specific environment/infrastructure domain to avoid unexpected problems that may imperil traffic safety and travel efficiency. Such capability awareness could be realized if the SAE level of autonomy is reported together with a specification of the infrastructure/environment domain needed for the smooth operation of that level of autonomy. In other words, the SAE level of autonomy could be reported with sub-levels within each level; sublevels corresponding to specified levels of infrastructure/environment are needed for the smooth operation of that sub-level of autonomy. Therefore, this paper proposes that the taxonomy should address not only the vehicle capability at each level of automation (as the current SAE taxonomy does) but rather the capabilities of the duo (the vehicle and its infrastructure/environment domain).

In this section of the paper, we propose a supplement that is intended to serve as an addition to the existing SAE taxonomy. We particularly focus on only Level 4. Therefore, the supplement is developed for Level 4 HADS only. Table 3, which presents a summary of the proposed supplement, shows the sublevels of Level 4 automation, the supporting infrastructure/environment associated with that sublevel, the features of the infrastructure/environment at each sublevel, and the capabilities of HADS corresponding to that level of the infrastructure/environment domain.

Table 3. The proposed supplement.

Sub-Level of Level 4	Characteristics of the Infrastructure/Environment Domain	Capabilities of the HADS
Level 4-A (Level 4 Vehicles on a Dedicated Guideway)	• Lanes are exclusive and fully controlled • Intelligent and complete infrastructure is accessible • Other road users such as pedestrians seldom occur	• Guideway following • Roadside parking
Level 4-B (Level 4 Vehicles on an Expressway)	• Reliable V2I and V2V communication is accessible • Surrounding vehicles are moving in the same direction • Other road users seldom occur. Wild animals may suddenly appear but at a relatively low frequency	• Trajectory planning • Mapping and localization at the lane level • Road language detection • Roadside parking

Table 3. *Cont.*

Sub-Level of Level 4	Characteristics of the Infrastructure/Environment Domain	Capabilities of the HADS
Level 4-C (Level 4 Vehicles on a Well-structured Road)	• Clear lane markers and complete traffic signals are accessible • Smart and communicable infrastructure may be inaccessible • A large number of other road users such as pedestrians and bicycles exist	• Trajectory planning • Mapping and localization at the lane level • Detect complex and multi-road language • Quick object and event detection and response • Parking at the parking lot and roadside
Level 4-D (Level 4 vehicles on a Limited-Structured Road)	• Road lane markers and traffic signs are incomplete or even unavailable • Intelligent infrastructure is usually inaccessible • The road may be covered by flood, ice, or dirt such that lane markers are invisible • Some wild animals, pedestrians, vehicles, and other road users may exist in the surroundings	• Trajectory planning and updating • Mapping and localization at centimeter-level • Detect complex and multi-road language • Quick object and event detection and response • Pass intersections without traffic signals • Parking
Level 4-E (Level 4 Vehicles in a Disorganized Area)	• The surroundings are constituted by huge crowds of people, bicycles, motors, and other road users • Space suitable for driving is usually limited • Assistance from nearby intelligent infrastructure is inaccessible	• Trajectory planning and updating • Mapping and localization more precise than centimeter-level • Detect complex and multi-road language • Object and event detection and response in a very dynamic and timely manner • Travel with the crowd at a low speed • Pass intersections without traffic signals • Parking

The supplement is motivated by the need for specifying the levels of autonomy together with physical or cyber infrastructure and environmental conditions. The Level 4 vehicle indicated in the table has capabilities consistent with any SAE Level 4 vehicle. It is important to note that this table is intended to provide a broad overview of possible scenarios and is not an exhaustive guideline for every possible circumstance. To bring the proposed levels of automation to life, consider the following examples: For Level 4-A, which necessitates a dedicated guideway for the normal operation of Level 4 HADS, a city bus running a regular route could be a practical example. The dedicated bus lane, equipped with traffic signals that prioritize the bus, can function as the dedicated guideway. The bus, outfitted with Level 4 automation, can operate safely within this controlled environment. For the case of Level 4-B, imagine an emergency vehicle such as an ambulance. These vehicles often benefit from traffic priority systems that can modify traffic signals to allow them to pass through intersections more quickly. However, the presence of these examples does not mean that all Level 4-A vehicles will be buses, or that all Level 4-B vehicles will be emergency vehicles. These are simply illustrative examples, intended to provide context and clarity for our proposed supplement.

The proposed supplement accounts for a key part of the ODD for automated systems classification. The ODD identifies specific limitations on the physical, cyber, environmental, geographical, and time-of-day restrictions and roadway characteristics. In the supplement, we focus only on the ODD elements that are within the control of the agency; therefore, weather-related elements such as ice, fog, heavy snow, or rain are excluded. The supplement is motivated by the need for specifying the levels of autonomy together with physical or cyber infrastructure and environmental conditions. Unlike physical and cyber conditions, weather-related conditions are not under the direct control of the stakeholders of the automated system, namely, the vehicle manufacturer, the agency, and the user.

The Level 4 vehicle indicated in the table has the following capabilities, consistent with any SAE Level 4 vehicle: (a) capable of adaptive cruise control, lane following, and lane transition; (b) capable of the basic detection of objects and events and the formulation of a response; (c) requires neither a conventional nor remote driver during route operation. It is assumed that there is no hardware or software system failure. The proposed sublevels of automation in the supplement are Levels 4-A, 4-B, 4-C, 4-D, and 4-E.

The designation of a specific sublevel of automation indicates its required operating condition. For example, Level 4-A depicts that a dedicated guideway is necessary to enable the normal operation of the Level 4 HADS. From Level 4-A to Level 4-E, there is a decrease in the quality requirement for the infrastructure/environment domain, for example, smart technology for road infrastructure.

5.1. Level 4-A (Level 4 Vehicles on a Dedicated Guideway)

Dedicated guideway conditions refer to lanes exclusively designed for AVs or areas with full access control. Examples of dedicated guideway conditions include automated bus-only lanes [93], Heathrow ULTra pods [36], and Tesla's AV-dedicated tracks [94]. Figure 2a shows autonomous pods at Heathrow Airport operating on a 3.9 km track that connects the terminals. These pods are only capable of unmanned driving on the dedicated tracks and are expected to park at pre-defined locations for passenger pick-up/drop-off. In Level 4-A, the entire automated system (vehicle and infrastructure/environment domain) is capable of guideway following and roadside parking. For the guideway-following task, the vehicle uses only the trajectory directed by the guideway and cooperates with enroute traffic lights, signs, and other roadside equipment. Such navigation may be supported by intelligent V2I communication. For the roadside parking task, the vehicle can perform temporary parking for passenger pick-up/drop-off or emergency and send out failure information when necessary.

5.2. Level 4-B (Level 4 Vehicles on an Expressway)

On a typical expressway, driving is characterized by reliable and communicable roadway infrastructure featuring high-quality, clear, and recognizable lane markers and legible signs, as shown in Figure 4. Surrounding vehicles move in the same direction, and platoons may be formed. When an AV joins a platoon, its automated driving system may use V2V communications to supplement its sensing capabilities. Similar to Level 4-A, Level 4-B automated systems (vehicle and infrastructure/environment domain) permit the AV to perform roadside parking. On an expressway, vehicles constitute the only road users; however, it is still possible for the AV to encounter anomalous environmental conditions including crossing animals or unruly pedestrians or two-wheelers. In encountering such situations, the AVs' HADS adjust their speed accordingly to maintain safe operations. In addition, in Level 4-B systems, HADS carry out the tasks of trajectory planning, mapping and localization, and road language detection. In planning a trajectory, the HADS decides the optimal route to the destination to minimize travel costs or travel distance using optimal route planning algorithms. For example, on a multi-lane expressway, optimal route planning could mean choosing the lane with the least congestion to minimize time, or selecting the lane that requires fewer lane changes to optimize fuel consumption and reduce wear. Furthermore, in cases where there are multiple routes available to reach the destination, such as with expressway junctions or interchanges, the HADS would determine the most efficient route based on current and predicted conditions. As a specific example, Magsino et al. [95] proposed an intelligent highway tollgate queue selector using fuzzy logic. Its purpose was to automatically select the most appropriate tollgate server for a vehicle to ensure the shortest waiting time. In performing the mapping and localization task, the AVs' HADS are expected to recognize their lane position accurately. This process might be supported by referential lane markers and frequency selective strips (FSSs) via radar detection or magnetic markers installed on the road surface [96]. Regarding road language detection, HADS are capable of detecting enroute traffic lights, signs, and other road language and interpreting the information from these facilities, often with the help of V2I communication.

Figure 4. An example of an expressway infrastructure–environment [97].

5.3. Level 4-C (Level 4 Vehicles on a Well-Structured Road)

This infrastructure/environment domain is characterized by a road with clear lane markers and complete traffic signals and signs. These include urban arterials and collectors. However, smart and communication infrastructure may be absent or inaccessible. Also, GPS signals may not be strong enough to provide localization support, particularly in tunnels or occluded areas. Also, there is a large volume of road user classes other than vehicles, particularly at intersection points in the corridor. Yet still, in Level 4-C, vehicles, and pedestrians use separate facilities. Figure 5 presents an example of a well-structured infrastructure/environment domain.

Figure 5. An example of a well-structured urban street infrastructure/environment domain in Shenzhen, China [98].

In operating safely at Level 4-C, the AVs' HADS detect objects and respond accurately and timely. During driving, HADS should identify all nearby road users using computer vision, predict their behavior, and take actions to avoid collision accordingly. This task is challenging because there are traffic lawbreakers who may either suddenly cross the road or violate traffic signals. Apart from timely object detection, HADS at this level have the capabilities of Level 4-B but offer these capabilities at a higher precision and with less external support. For the trajectory planning task, the HADS of this level decide the optimal route to the destination with considerations of real-time road traffic. The decision can be updated using enroute real-time traffic information to avoid potential congestion [99–101]. For the mapping and localization task, the vehicle should be able to recognize its location with at least lane-level accuracy using its on-board vision system. Compared to the Level 4-B expressway level, the localization task is much more challenging as the road may be narrow on minor streets. Also, as supporting infrastructure such as magnetic lane markers may be absent, the HADS detect lanes solely using their onboard sensing features. About the capability to detect road language, the HADS identify not only permanent traffic signals and signs but also temporary signals including traffic police hand directions and short-term detour signs at construction work zones. The AV is able to pass through an intersection safely with guidance from traffic signals and signs.

5.4. Level 4-D (Level 4 Vehicles on a Limited-Structured Road)

At level 4-D, the infrastructure/environment domain is characterized by undeveloped off-road conditions/limited-structured roads. Examples of limited-structured roads include rural gravel roads (Figure 6), dirt tracks, intersections and roads without traffic signs, deserts, and frozen lakes. Road lane markers, traffic signs, and intelligent infrastructure are incomplete or even unavailable in these areas. The off-road nature of the driving environment and lack of infrastructure severely impair the capability of the Level 4 vehicle from realizing the full potential capabilities of SAE-defined automation at Level 4. In this case, the entire system (the Level 4 vehicle and its infrastructure/environment domain) is described as Level 4-D. In such environments, HADS plan their trajectory with little or no support from surrounding infrastructure and do this largely using onboard sensors to delineate the geometry of the road space, including boundaries of the pavement and vegetation, and to recognize pedestrians, bicycles, and other vehicles that may intrude their trajectory and adjust accordingly. At an unsignalized intersection, the HADS use their sensors to establish or predict the movements of other road users and to negotiate with them in order to pass the intersection safely.

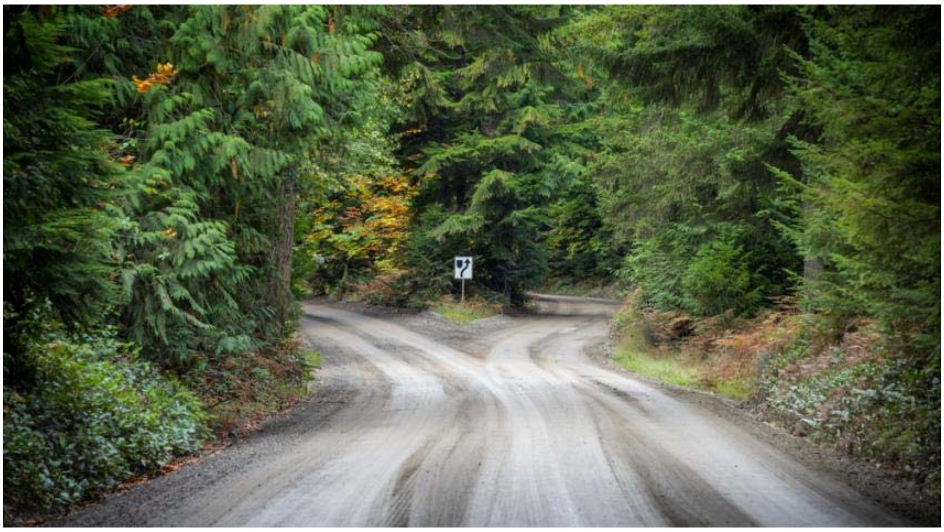


Figure 6. An example of a limited-structured road infrastructure [102].

5.5. Level 4-E (Level 4 Vehicles in a Disorganized Area)

The most challenging infrastructure/environment domain for an automated driving system is one where there is an absolute lack of structure and organization in the way the road space is used (Figure 7). Often, such road spaces are dominated by pedestrians and two-wheelers without clear paths milling around the road space in a seemingly random manner [51,103]. Road lane markers and traffic signs are often not visible or non-existent. Examples include pedestrian zones in urban areas, outdoor markets, sports squares, and crowded plazas. The highly dynamic nature of this infrastructure/environment domain severely impairs the capability of a Level 4 vehicle to operate in the manner expected of an SAE Level 4 vehicle. The vehicle is expected to travel within the crowd at low speed and to continuously assess the proximity of the surroundings even though the view of vehicles might be partially blocked by moving humans. Assistance from nearby intelligent infrastructure is neither expected nor provided due to pedestrian overcrowding and the highly dynamic nature of the environment.

Figure 7. An example of driving at the level of a disorganized area [104].

As Figure 7 suggests, driving in disorganized environments can be even more complex compared to limited-structured roads. In the task of trajectory planning in such an environment, HADS should be capable of not only evaluating the spatial limitations and choosing the safest and most route for navigating through the crowd (e.g., the route with the lowest density of pedestrians) but also responding to changes in that environment in a timelier manner compared to the HADS of Level 4-D (Level 4 Vehicles on a Limited-Structured Road).

To facilitate their movement in the crowd, the HADS construct a real-time profile of the moving crowd, estimate their position in the crowd, and calculate relative distances to surrounding objects (individual pedestrians and other moving and stationary features) using their onboard sensors. The surrounding traffic consists of automobiles, two-wheelers, and people; therefore, mapping and localization requirements are high. The HADS are able to recognize the type, size, locations, directions, and intentions of other road users. Extremely dynamic disorganized scenes also raise challenges in driving strategies: the vehicle is supposed to travel through the crowd at a safe speed. The vehicle may need to perform stop-and-go behavior frequently, detour to bypass pedestrians, or inch forward to make space in a dense crowd [105]. At unsignalized intersections, similar to Level 4-D, Level 4-E's HADS predict the actions of other road users and negotiate with them to pass the intersection safely. It is worth mentioning the parking task for this proposed level: the vehicle should be able to park itself on roadside or parking spots inside the crowd even when the boundaries of the parking place are not visible. Compared to other levels, parking is more demanding at this level because the HADS may be expected to be able to find a parking spot in an area that is crowded with milling pedestrians.

6. Discussion

6.1. The Proposed Supplement

Section 5 documents the establishment of a five-level classification supplement for Level 4 SAE automation, through an examination of the different infrastructure/environment domains and the capabilities of Level 4 HADS under those domains. It is noteworthy that the necessary capabilities for HADS differ across various levels. Some levels may not necessitate certain advanced capabilities as the driving scenarios at that level do not require such capabilities. On the other hand, certain levels may require more demanding capabilities for specific functions owing to the nature of the driving conditions at that particular level. For example, trajectory planning functions are not needed for Level 4-A because the HADS are generally guided by the guideway. The requirement of object recognition is relatively low for HADS of Level 4-A and Level 4-B because strict access control is applied for their driving scenes and road users other than vehicles are rare. In contrast, a fast and accurate vision system is necessary for HADS of Level 4-E due to the existence of a large number of pedestrians.

Each level of the proposed supplement indicates a mandatory minimum rather than maximum capabilities for HADS operations at that level. It is not possible to specify a complete set of features for each level of HADS since ADS-dedicated vehicles have not been extensively implemented. The details of each sublevel may be expanded in the future to reflect technology-driven advancements, and new sublevels may be added in the future.

6.2. The Future of HADS and Infrastructure Development

Using the proposed classification supplement, it is easier to characterize the capability of automated driving systems in a wide range of operating conditions. Accessibility to cyber infrastructure, the complexity of driving algorithms, and the domain of driving tasks vary across the different infrastructure conditions. For example, accurate digital maps with high definition which assist localization and detection are not available for all locations, particularly those in rural areas. On the other hand, the traffic conditions in urban areas are significantly more complex compared to rural areas. As a result, autonomous vehicles rely not only on the on-board sensors for navigation, but also heavily depend on infrastructure [29]. In this regard, certain studies focus on researching the optimal deployment of roadside units (RSUs) [106], the use of fog computing for efficient data exchange [68,88], and the identification of crucial intersections [107], aiming to expedite the integration of autonomous vehicles in urban environments.

In recognition of the nature of different levels of HADS, we provide Figure 8 to visualize the concept of potential differences among levels of HADS regarding their dependence on intelligent infrastructure and in-vehicle artificial intelligence (AI) systems. The AI systems carried by vehicles are expected to perceive the surrounding driving environment and make informed decisions in a timely manner. This is particularly critical for HADS of Level 4-D and Level 4-E, where the vehicle needs to drive in highly dynamic and complicated situations and operate in areas with inadequate supporting infrastructure and/or challenging environments. A powerful onboard AI system would make the AV capable of navigating challenging and diverse driving scenarios. However, a flexible and robust system is still far from reality. Ref. [108] reviewed state-of-the-art computer vision models used in automated driving, and pointed out that existing models are still inferior to human perception and reasoning. Ref. [109] suggested that such decision-making algorithms lack extensive real-world tests.

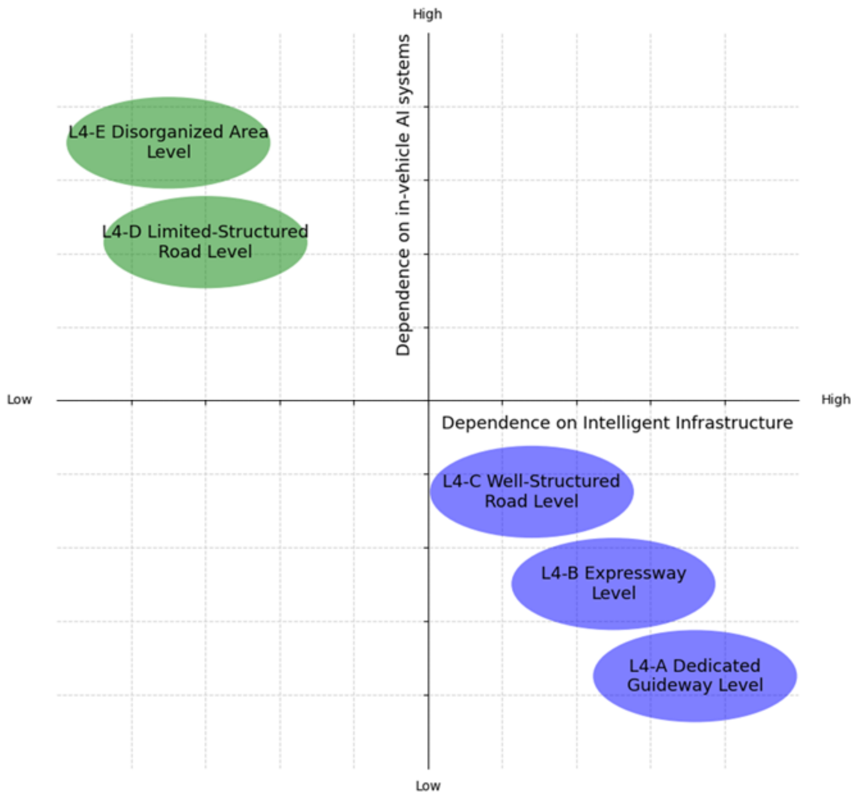


Figure 8. Visualization of differences among different levels of HADS.

Infrastructure plays a crucial role in certain levels of HADS such as Level 4-A, Level 4-B, and Level 4-C through providing reliable traffic information, guiding the vehicle through the high-quality network, and even sending advisory instructions to AVs. These levels of AVs are easier to reach compared to other levels since existing research has studied how infrastructure can be built into applications [93,110–112]. For example, traffic lights were used to assign priority when more than two vehicles enter an intersection in UK's Autodrive Project [111], and traffic accidents were avoided. Ye and Yamamoto also investigated the positive impact of setting dedicated lanes for AVs on traffic flow throughput [112]. These studies pointed out that superior system efficiency, a lower fatality rate, and improved fuel economy could be expected when infrastructure is made to play a key role in automated driving, which is likely to be the case for HADS of Level 4-A and Level 4-B.

7. Concluding Remarks

This study is based on the premise that high-level AVs, in the future, will not only be deployed in areas where they have been tested to exhibit full capabilities of Level 4 vehicle automation, but also in areas where the infrastructure and environment may not be so conducive to support such full capabilities. In that case, the existing SAE taxonomy may be inadequate for characterizing the capability of these systems and their ODD. To throw more light and stimulate discussion on this issue, this study proposed a five-level (SAE sublevels) classification framework to serve as a supplement to the current SAE taxonomy for highly automated driving systems (HADS). The study identified features of the infrastructure and environment as well as the minimum capabilities of HADS at each sub-level. In contrast to a school of thought which postulates that the capabilities of HADS should be viewed only against the functions of human drivers in the automated system, the proposed supplement emphasizes the inclusion of the infrastructure and corresponding driving environment in the classification of automated systems. The study recognizes that different capability levels of high-level automated driving are realized under different infrastructure and environmental domains, and therefore, the classification should be based not only on the capability of the vehicle (and thus, the functions of the human driver) but should be based on both the vehicle capability and the infrastructure–environment within which it operates. Through examining the significance of different infrastructure conditions, we proposed five sublevels of Level 4 HADS: Level 4-A (Dedicated Guideway Level), Level 4-B (Expressway Level), Level 4-C (Well-Structured Road Level), Level 4-D (Limited-Structured Road Level), and Level 4-E (Disorganized Area Level). From Level 4-A to Level 4-E, (a) the environment becomes less orderly and increasingly chaotic, and (b) support from nearby intelligent infrastructure decreases, and therefore, onboard navigation systems play an increasingly greater role.

The proposed supplement is descriptive rather than normative. The supplement will clarify the realistic capabilities of the ADS, the condition of use, and the requisite supporting infrastructure. AV developers and manufacturers are herein provided with a basis upon which they can provide the more accurate specification of the capabilities of their products from the ODD perspective; that way, user expectations regarding travel efficiency and safety of the ADS at different locations and environments can be clearer. In other words, the supplement will help clear up any confusion regarding the capabilities and limitations of AVs.

In future studies, work can be carried out to further improve certain aspects of the proposed supplement. First, to enhance further the applicability of the supplement, future efforts could examine the classification of infrastructure intelligence, and incorporate the dynamics of traffic scenes and road conditions. Second, the framework is based on the current state of the AV market, where the automation level is close to the SAE's Level 3. However, it can be envisioned that various levels of current roadway design features including lane and shoulder width, horizontal and vertical alignment, and cross-section will need to be evaluated to facilitate AV traffic operations. Therefore, with the rapid development of automated driving technologies, potential adjustments could be applied to improve the feasibility and compatibility of our proposed supplement. In the future, a well-defined scientific framework that addresses the issue of varying capabilities required at distinct levels of HADS is imperative. Such a framework should provide a normative and quantitative specification of the essential capabilities for the different levels of HADS while additionally clarifying the respective degree of dependence on intelligent infrastructure or artificial intelligence. This paper's Figure 8 may serve as a guiding framework for such discussions. Third, using real-world (and possibly in-service) data, quantitative comparisons between the capabilities of different HADS could be carried out in order to avoid descriptive comparisons as done in this paper.

Overall, the supplement presented in this paper is expected to be beneficial in clarifying potential sublevels of Level 4 automated driving systems and enhancing the SAE's taxonomical classification. Reference to the proposed supplement is expected to benefit prospective AV consumers and vehicle manufacturers in terms of setting a clear and readily understandable bar of performance expectations in different environments and infrastructure settings. In addition, government road agencies will be placed in a better position to justify investments in infrastructure geared towards improving their infrastructure to support the coming age of driving automation.

Author Contributions: Conceptualization, S.C., S.Z., T.C., Z.H., Y.C. and S.L.; formal analysis, S.C., S.Z. and Y.C.; funding acquisition, S.L.; methodology, Y.C., S.Z. and S.C.; project administration, S.C. and S.L.; supervision, S.C. and S.L.; validation, T.C. and Z.H.; visualization, S.Z.; writing—original draft, S.C., S.Z., T.C., Z.H., Y.C. and S.L.; writing—review and editing, S.C., T.C., Y.C. and S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abe, R. Introducing autonomous buses and taxis: Quantifying the potential benefits in Japanese transportation systems. *Transp. Res. Part A Policy Pract.* **2019**, *126*, 94–113. [CrossRef]
2. Feng, S.; Sun, H.; Yan, X.; Zhu, H.; Zou, Z.; Shen, S.; Liu, H.X. Dense reinforcement learning for safety validation of autonomous vehicles. *Nature* **2023**, *615*, 620–627. [CrossRef]
3. Kolarova, V.; Steck, F.; Bahamonde-Birke, F.J. Assessing the effect of autonomous driving on value of travel time savings: A comparison between current and future preferences. *Transp. Res. Part A Policy Pract.* **2019**, *129*, 155–169. [CrossRef]
4. USDOT. Automated Vehicles 3.0: Preparing for the Future of Transportation. 2018. Available online: https://www.transportation.gov/sites/dot.gov/files/docs/policy-initiatives/automated-vehicles/320711/preparing-future-transportation-automated-vehicle-30.pdf (accessed on 19 June 2023).
5. Badue, C.; Guidolini, R.; Carneiro, R.V.; Azevedo, P.; Cardoso, V.B.; Forechi, A.; Jesus, L.; Berriel, R.; Paixão, T.M.; Mutz, F.; et al. Self-driving cars: A survey. *Expert Syst. Appl.* **2021**, *165*, 113816. [CrossRef]
6. Bagloee, S.A.; Tavana, M.; Asadi, M.; Oliver, T. Autonomous vehicles: Challenges, opportunities, and future implications for transportation policies. *J. Mod. Transp.* **2016**, *24*, 284–303. [CrossRef]
7. Kurniawan, J.H.; Chng, S.; Cheah, L. The Social Acceptance of Autonomous Vehicles. *IEEE Potentials* **2021**, *40*, 39–44. [CrossRef]
8. USDOT. Automated Vehicles Comprehensive Plan. 2021. Available online: https://www.transportation.gov/av/avcp (accessed on 19 June 2023).
9. FHWA CARMA. Overview. 2019. Available online: https://highways.dot.gov/research/research-programs/operations/carma-overview (accessed on 19 June 2023).
10. AASHTO. Connecting on CAVs. 2021. Available online: https://cav.transportation.org/wp-content/uploads/sites/61/2021/11/AASHTO-CAV-Breakfast-Meeting-10-21-2021.pdf (accessed on 19 June 2023).
11. Bainbridge, L. Ironies of Automation. In *Analysis, Design and Evaluation of Man-Machine Systems*; Elsevier: Amsterdam, The Netherlands, 1983; pp. 129–135. ISBN 978-0-08-029348-6.
12. Rudolph, G.; Uwe, V. Three Sensor Types Drive Autonomous Vehicles. 2017. Available online: https://www.fierceelectronics.com/components/three-sensor-types-drive-autonomous-vehicles (accessed on 19 June 2023).
13. SAE. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. 2018. Available online: https://www.sae.org/standards/content/j3016_201806/ (accessed on 19 June 2023).
14. Darbha, S.; Konduri, S.; Pagilla, P.R. Benefits of V2V Communication for Autonomous and Connected Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 1954–1963. [CrossRef]
15. Wuthishuwong, C.; Traechtler, A. Vehicle to infrastructure based safe trajectory planning for Autonomous Intersection Management. In Proceedings of the 2013 13th International Conference on ITS Telecommunications (ITST), Tampere, Finland, 5–7 November 2013; IEEE: Tampere, Finland, 2013; pp. 175–180.
16. Fraedrich, E.; Heinrichs, D.; Bahamonde-Birke, F.J.; Cyganski, R. Autonomous driving, the built environment and policy implications. *Transp. Res. Part A Policy Pract.* **2019**, *122*, 162–172. [CrossRef]
17. Storsæter, A.D.; Pitera, K.; McCormack, E.D. The automated driver as a new road user. *Transp. Rev.* **2021**, *41*, 533–555. [CrossRef]
18. OICA. Automated Driving Definition for Levels of Automation 2014. Available online: https://www.oica.net/wp-content/uploads/WP-29-162-20-OICA-automated-driving.pdf (accessed on 19 June 2023).
19. Gasser, T.M.; Westhoff, D. BASt-study: Definitions of Automation and Legal Issues in Germany. In Proceedings of the 2012 Road Vehicle Automation Workshop. Automation Workshop. 2012. Available online: https://onlinepubs.trb.org/onlinepubs/conferences/2012/Automation/presentations/Gasser.pdf (accessed on 19 June 2023).
20. NHTSA. Automated Driving Systems: A Vision for Safety. 2017. Available online: https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf (accessed on 19 June 2023).
21. Geisler, T.; Shi, E. Taxonomies of Connected, Cooperative and Automated Mobility. In Proceedings of the 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 4–9 June 2022; IEEE: Aachen, Germany, 2022; pp. 1517–1524.
22. Rana, M.; Hossain, K. Connected and Autonomous Vehicles and Infrastructures: A Literature Review. *Int. J. Pavement Res. Technol.* **2021**, *16*, 264–284. [CrossRef]
23. Stayton, E.; Stilgoe, J. It's Time to Rethink Levels of Automation for Self-Driving Vehicles [Opinion]. *IEEE Technol. Soc. Mag.* **2020**, *39*, 13–19. [CrossRef]
24. Steckhan, L.; Spiessl, W.; Quetschlich, N.; Bengler, K. Beyond SAE J3016: New Design Spaces for Human-Centered Driving Automation. In *HCI in Mobility, Transport, and Automotive Systems*; Krömker, H., Ed.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2022; Volume 13335, pp. 416–434. ISBN 978-3-031-04986-6.
25. Tengilimoglu, O.; Carsten, O.; Wadud, Z. Implications of automated vehicles for physical road environment: A comprehensive review. *Transp. Res. Part E Logist. Transp. Rev.* **2023**, *169*, 102989. [CrossRef]
26. Tengilimoglu, O.; Carsten, O.; Wadud, Z. Infrastructure requirements for the safe operation of automated driving: Opinions from experts and stakeholders. *Transp. Policy* **2023**, *133*, 209–222. [CrossRef]
27. Huang, Z.; Zhu, X.; Lin, Y.; Xu, L.; Mao, Y. A novel WIFI-oriented RSSI signal processing method for tracking low-speed pedestrians. In Proceedings of the 2019 5th International Conference on Transportation Information and Safety (ICTIS), Liverpool, UK, 14–17 July 2019; pp. 1018–1023.

28. Lin, Y.; Huang, Z.; Wu, P.; Xu, L. RSSI positioning method of vehicles in tunnels based on semi-supervised extreme learning machine. *J. Traffic Transp. Eng.* **2021**, *21*, 243–255.
29. Huang, Z.; Chen, S.; Pian, Y.; Sheng, Z.; Ahn, S.; Noyce, D.A. CV2X-LOCA: Roadside Unit-Enabled Cooperative Localization Framework for Autonomous Vehicles. *arXiv* **2023**, arXiv:230400676.
30. Anastasiadou, K.; Vougias, S. "Smart" or "sustainably smart" urban road networks? The most important commercial street in Thessaloniki as a case study. *Transp. Policy* **2019**, *82*, 18–25. [CrossRef]
31. Joshi, S.; Bailey, A.; Datta, A. On the move? Exploring constraints to accessing urban mobility infrastructures. *Transp. Policy* **2021**, *102*, 61–74. [CrossRef]
32. Hatzenbühler, J.; Cats, O.; Jenelius, E. Transitioning towards the deployment of line-based autonomous buses: Consequences for service frequency and vehicle capacity. *Transp. Res. Part A Policy Pract.* **2020**, *138*, 491–507. [CrossRef]
33. AASHTO. AASHTO Policy Letters Focus On The Needs Of Autonomous Vehicles. 2018. Available online: https://aashtojournal.org/2018/11/30/aashto-policy-letters-focus-on-the-needs-of-autonomous-vehicles/ (accessed on 19 June 2023).
34. McAslan, D.; Gabriele, M.; Miller, T.R. Planning and policy directions for autonomous vehicles in metropolitan planning organizations (MPOs) in the United States. *J. Urban Technol.* **2021**, *28*, 175–201. [CrossRef]
35. CBS News. Elon Musk and Gayle King Test Drive the Tunnel He Hopes Will Solve L.A. Traffic 2018. Available online: https://www.cbsnews.com/news/elon-musk-inside-boring-company-tunnel-with-gayle-king-spacex/ (accessed on 19 June 2023).
36. University of Bristol. Heathrow Pods-Personal Rapid Transit | Research | University of Bristol. 2016. Available online: http://www.bristol.ac.uk/research/impact/heathrow-personal-rapid-transport/ (accessed on 19 June 2023).
37. Jacksonville, Fla. Ultimate Urban Circulator. Available online: https://u2c.jtafla.com/ (accessed on 19 June 2023).
38. Darrell, E. Self-Driving Mercedes-Benz Bus Takes a Milestone 12-Mile Trip. 2016. Available online: https://techcrunch.com/2016/07/18/self-driving-mercedes-benz-bus-takes-a-milestone-12-mile-trip/ (accessed on 19 June 2023).
39. Rachel, H. A Ride On Heathrow's Self-Driving Pods. 2018. Available online: https://londonist.com/2014/09/a-ride-on-heathrows-self-driving-pods (accessed on 19 June 2023).
40. Ran, B.; Cheng, Y.; Li, S.; Li, H.; Parker, S. Classification of Roadway Infrastructure and Collaborative Automated Driving System. *SAE Int. J. Connect. Autom. Veh.* **2023**, *6*, 12-06-04-0026. [CrossRef]
41. Saeed, T.U.; Alabi, B.N.T.; Labi, S. Preparing Road Infrastructure to Accommodate Connected and Automated Vehicles: System-Level Perspective. *J. Infrastruct. Syst.* **2021**, *27*, 6020003. [CrossRef]
42. Soteropoulos, A.; Mitteregger, M.; Berger, M.; Zwirchmayr, J. Automated drivability: Toward an assessment of the spatial deployment of level 4 automated vehicles. *Transp. Res. Part A Policy Pract.* **2020**, *136*, 64–84. [CrossRef]
43. Manivasakan, H.; Kalra, R.; O'Hern, S.; Fang, Y.; Xi, Y.; Zheng, N. Infrastructure requirement for autonomous vehicle integration for future urban and suburban roads—Current practice and a case study of Melbourne, Australia. *Transp. Res. Part A Policy Pract.* **2021**, *152*, 36–53. [CrossRef]
44. Edwards, S.; Hill, G.; Goodman, P.; Blythe, P.; Mitchell, P.; Huebner, Y. Quantifying the impact of a real world cooperative-ITS deployment across multiple cities. *Transp. Res. Part A Policy Pract.* **2018**, *115*, 102–113. [CrossRef]
45. Huang, Z.; Xu, L.; Lin, Y. Multi-stage pedestrian positioning using filtered WiFi scanner data in an urban road environment. *Sensors* **2020**, *20*, 3259. [CrossRef] [PubMed]
46. Shladover, S.E.; Desoer, C.A.; Hedrick, J.K.; Tomizuka, M.; Walrand, J.; Zhang, W.-B.; McMahon, D.H.; Peng, H.; Sheikholeslam, S.; McKeown, N. Automated vehicle control developments in the PATH program. *IEEE Trans. Veh. Technol.* **1991**, *40*, 114–130. [CrossRef]
47. Fernandes, P.; Nunes, U. Platooning With IVC-Enabled Autonomous Vehicles: Strategies to Mitigate Communication Delays, Improve Safety and Traffic Flow. *IEEE Trans. Intell. Transp. Syst.* **2012**, *13*, 91–106. [CrossRef]
48. Jia, D.; Ngoduy, D. Enhanced cooperative car-following traffic model with the combination of V2V and V2I communication. *Transp. Res. Part B Methodol.* **2016**, *90*, 172–191. [CrossRef]
49. CBingsights 40+ Corporations Working on Autonomous Vehicles. 2021. Available online: https://www.cbinsights.com/research/autonomous-driverless-vehicles-corporations-list/ (accessed on 19 June 2023).
50. China Automotive News GAC Group Launches Aion Level 4 Self-Driving Version. 2021. Available online: https://autonews.gasgoo.com/ICV/70016362.html (accessed on 19 June 2023).
51. Cai, P.; Luo, Y.; Saxena, A.; Hsu, D.; Lee, W.S. LeTS-Drive: Driving in a Crowd by Learning from Tree Search. *arXiv* **2019**, arXiv:1905.12197.
52. Chen, S.; Saeed, T.U.; Labi, S. Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. *Anal. Methods Accid. Res.* **2017**, *16*, 75–89. [CrossRef]
53. Chen, S.; Leng, Y.; Labi, S. A deep learning algorithm for simulating autonomous driving considering prior knowledge and temporal information. *Comput.-Aided Civ. Infrastruct. Eng.* **2020**, *35*, 305–321. [CrossRef]
54. Gomes, L. Urban Jungle a Tough Challenge for Google's Autonomous Cars. *MIT Technol. Rev.* **2014**. Available online: https://www.technologyreview.com/2014/07/24/171973/urban-jungle-a-tough-challenge-for-googles-autonomous-cars/ (accessed on 19 June 2023).
55. Knight, W. Can This Man Make AI More Human? 2015. Available online: https://www.technologyreview.com/2015/12/17/164285/can-this-man-make-ai-more-human/ (accessed on 19 June 2023).

56. Silberg, G.; Wallace, R. Selfdriving Cars: The Next Revolution. 2012. Available online: http://www.cargroup.org/wp-content/uploads/2017/02/Self_driving-cars-The-next-revolution.pdf (accessed on 19 June 2023).
57. Sheng, Z.; Huang, Z.; Chen, S. EPG-MGCN: Ego-Planning Guided Multi-Graph Convolutional Network for Heterogeneous Agent Trajectory Prediction. *arXiv* **2023**, arXiv:230317027.
58. Waymo Safety Report. 2021. Available online: https://downloads.ctfassets.net/sv23gofxcuiz/4gZ7ZUxd4SRj1D1W6z3rpR/2ea16814cdb42f9e8eb34cae4f30b35d/2021-03-waymo-safety-report.pdf (accessed on 19 June 2023).
59. Jalihal, D.S.A.; Ravinder, K. Traffic Characteristics of India. In Proceedings of the Eastern Asia Society for Transportation Studie, 2015; p. 5. Available online: https://www.researchgate.net/publication/237535142_TRAFFIC_CHARACTERISTICS_OF_INDIA (accessed on 19 June 2023).
60. Li, J.; Xu, L.; Li, R.; Wu, P.; Huang, Z. Deep spatial-temporal bi-directional residual optimisation based on tensor decomposition for traffic data imputation on urban road network. *Appl. Intell.* **2022**, *52*, 11363–11381. [CrossRef]
61. Kotseruba, I.; Rasouli, A.; Tsotsos, J.K. Joint Attention in Autonomous Driving (JAAD). *arXiv* **2016**, arXiv:1609.04741.
62. McAslan, D.; Najar Arevalo, F.; King, D.A.; Miller, T.R. Pilot project purgatory? Assessing automated vehicle pilot projects in U.S. cities. *Humanit. Soc. Sci. Commun.* **2021**, *8*, 325. [CrossRef]
63. Liu, X.; Liu, Y.; Chen, Y.; Wang, L.; Lu, Z. Reinforcement Learning in V2I Communication Assisted Autonomous Driving. In Proceedings of the ICC 2020–2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; IEEE: Dublin, Ireland, 2020; pp. 1–6.
64. Milanes, V.; Villagra, J.; Godoy, J.; Simo, J.; Perez, J.; Onieva, E. An Intelligent V2I-Based Traffic Management System. *IEEE Trans. Intell. Transp. Syst.* **2012**, *13*, 49–58. [CrossRef]
65. Outay, F.; Mengash, H.A.; Adnan, M. Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: Recent advances and challenges. *Transp. Res. Part A Policy Pract.* **2020**, *141*, 116–129. [CrossRef]
66. Xu, R.; Huang, Z.; Chen, S.; Li, J.; Wu, P.; Lin, Y. Wi-CL: Low-Cost WiFi-Based Detection System for Nonmotorized Traffic Travel Mode Classification. *J. Adv. Transp.* **2023**, *2023*, 1033717. [CrossRef]
67. Chen, S.; Dong, J.; Ha, P.; Li, Y.; Labi, S. Graph neural network and reinforcement learning for multi-agent cooperative control of connected autonomous vehicles. *Comput. Aided Civ. Infrastruct. Eng.* **2021**, *36*, 838–857. [CrossRef]
68. Du, R.; Chen, S.; Dong, J.; Ha, P.Y.J.; Labi, S. GAQ-EBkSP: A DRL-based Urban Traffic Dynamic Rerouting Framework using Fog-Cloud Architecture. In Proceedings of the 2021 IEEE International Smart Cities Conference (ISC2), Manchester, UK, 7–10 September 2021; IEEE: Manchester, UK, 2021; pp. 1–7.
69. Ha, P.; Chen, S.; Du, R.; Dong, J.; Li, Y.; Labi, S. Vehicle Connectivity and Automation: A Sibling Relationship. *Front. Built Environ.* **2020**, *6*, 590036. [CrossRef]
70. Huang, L.; Guo, H.; Zhang, R.; Wang, H.; Wu, J. Capturing Drivers' Lane Changing Behaviors on Operational Level by Data Driven Methods. *IEEE Access* **2018**, *6*, 57497–57506. [CrossRef]
71. Wu, P.; Huang, Z.; Pian, Y.; Xu, L.; Li, J.; Chen, K. A Combined Deep Learning Method with Attention-Based LSTM Model for Short-Term Traffic Speed Forecasting. *J. Adv. Transp.* **2020**, *2020*, 8863724. [CrossRef]
72. Anaya, J.J.; Merdrignac, P.; Shagdar, O.; Nashashibi, F.; Naranjo, J.E. Vehicle to pedestrian communications for protection of vulnerable road users. In Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings, Dearborn, MI, USA, 8–11 June 2014; IEEE: Dearborn, MI, USA, 2014; pp. 1037–1042.
73. Sugimoto, C.; Nakamura, Y.; Hashimoto, T. Prototype of pedestrian-to-vehicle communication system for the prevention of pedestrian accidents using both 3G wireless and WLAN communication. In Proceedings of the 2008 3rd International Symposium on Wireless Pervasive Computing, Santorini, Greece, 7–9 May 2008; IEEE: Santorini, Greece, 2008; pp. 764–767.
74. Wu, X.; Miucic, R.; Yang, S.; Al-Stouhi, S.; Misener, J.; Bai, S.; Chan, W. Cars Talk to Phones: A DSRC Based Vehicle-Pedestrian Safety System. In Proceedings of the 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall), Vancouver, BC, Canada, 14–17 September 2014; IEEE: Vancouver, BC, Canada, 2014; pp. 1–7.
75. Krok, A. Nikola Motor Company: The EV Startup with the Worst, Most Obvious Name Ever. 2016. Available online: https://www.cnet.com/roadshow/news/nikola-motor-company-the-ev-startup-with-the-worst-most-obvious-name-ever/ (accessed on 19 June 2023).
76. Welch, T.F.; Widita, A. Big data in public transportation: A review of sources and methods. *Transp. Rev.* **2019**, *39*, 795–818. [CrossRef]
77. Zhang, M.; Zhao, P.; Qiao, S. Smartness-induced transport inequality: Privacy concern, lacking knowledge of smartphone use and unequal access to transport information. *Transp. Policy* **2020**, *99*, 175–185. [CrossRef]
78. Huang, Z.; Xu, L.; Lin, Y.; Wu, P.; Feng, B. Citywide metro-to-bus transfer behavior identification based on combined data from smart cards and GPS. *Appl. Sci.* **2019**, *9*, 3597. [CrossRef]
79. Fleyeh, H.; Davami, E. Eigen-based traffic sign recognition. *IET Intell. Transp. Syst.* **2011**, *5*, 190. [CrossRef]
80. Taamneh, M.; Alkheder, S. Traffic sign perception among Jordanian drivers: An evaluation study. *Transp. Policy* **2018**, *66*, 17–29. [CrossRef]
81. Dong, J.; Chen, S.; Li, Y.; Du, R.; Steinfeld, A.; Labi, S. Space-weighted information fusion using deep reinforcement learning: The context of tactical control of lane-changing autonomous vehicles and connectivity range assessment. *Transp. Res. Part C Emerg. Technol.* **2021**, *128*, 103192. [CrossRef]

82. Jin, J.; Fu, K.; Zhang, C. Traffic Sign Recognition with Hinge Loss Trained Convolutional Neural Networks. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 1991–2000. [CrossRef]
83. Qian, R.; Zhang, B.; Yue, Y.; Wang, Z.; Coenen, F. Robust chinese traffic sign detection and recognition with deep convolutional neural network. In Proceedings of the 2015 11th International Conference on Natural Computation (ICNC), Zhangjiajie, China, 15–17 August 2015; IEEE: Zhangjiajie, China, 2015; pp. 791–796.
84. Zhu, Z.; Liang, D.; Zhang, S.; Huang, X.; Li, B.; Hu, S. Traffic-Sign Detection and Classification in the Wild. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Las Vegas, NV, USA, 2016; pp. 2110–2118.
85. Levinson, J.; Askeland, J.; Becker, J.; Dolson, J.; Held, D.; Kammel, S.; Kolter, J.Z.; Langer, D.; Pink, O.; Pratt, V.; et al. Towards fully autonomous driving: Systems and algorithms. In Proceedings of the 2011 IEEE Intelligent Vehicles Symposium (IV), Baden-Baden, Germany, 5–9 June 2011; IEEE: Baden-Baden, Germany, 2011; pp. 163–168.
86. Dong, J.; Chen, S.; Zong, S.; Chen, T.; Labi, S. Image transformer for explainable autonomous driving system. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021; IEEE: Indianapolis, IN, USA, 2021; pp. 2732–2737.
87. Fu, M.-Y.; Huang, Y.-S. A survey of traffic sign recognition. In Proceedings of the 2010 International Conference on Wavelet Analysis and Pattern Recognition, Qingdao, China, 11–14 July 2010; IEEE: Qingdao, China, 2010; pp. 119–124.
88. Ha, P.; Chen, S.; Du, R.; Labi, S. Scalable Traffic Signal Controls Using Fog-Cloud Based Multiagent Reinforcement Learning. *Computers* **2022**, *11*, 38. [CrossRef]
89. Franke, U.; Gavrila, D.; Gorzig, S.; Lindner, F.; Puetzold, F.; Wohler, C. Autonomous driving goes downtown. *IEEE Intell. Syst.* **1998**, *13*, 40–48. [CrossRef]
90. California Department of Transportation California Manual for Setting Speed Limits, Sacramento, CA, Retrieved August 6, 2021. 2014. Available online: https://cal.streetsblog.org/wp-content/uploads/sites/13/2019/07/california-manual-for-setting-speed-limits.pdf (accessed on 19 June 2023).
91. Du, R.; Chen, S.; Li, Y.; Alinizzi, M.; Labi, S. A Framework for Lane-Change Maneuvers of Connected Autonomous Vehicles in a Mixed-Traffic Environment. *Electronics* **2022**, *11*, 1350. [CrossRef]
92. McCall, J.C.; Trivedi, M.M. Video-Based Lane Estimation and Tracking for Driver Assistance: Survey, System, and Evaluation. *IEEE Trans. Intell. Transp. Syst.* **2006**, *7*, 20–37. [CrossRef]
93. Mercedes-Benz Mercedes Created the Bus of the Future—and It Looks Awesome. Business Insider. 2017. Available online: https://www.businessinsider.com/mercedes-future-bus-city-pilot-photos-features-2017-6?international=true&r=US&IR=T (accessed on 19 June 2023).
94. HTAuto Elon Musk Is Charged about Prospects of Self-Driving EVs in Tunnels under Cities. 2020. Available online: https://auto.hindustantimes.com/auto/news/elon-musk-is-charged-about-prospects-of-self-driving-evs-in-tunnels-under-cities-41600227254920.html (accessed on 19 June 2023).
95. Magsino, E.R.; Ho, I.W.H. An intelligent highway tollgate queue selector for improving server utilization and vehicle waiting time. In Proceedings of the 2016 IEEE Region 10 Symposium (TENSYMP), Bali, Indonesia, 9–11 May 2016; IEEE: Bali, Indonesia, 2016; pp. 271–276.
96. Byun, Y.-S.; Jeong, R.-G.; Kang, S.-W. Vehicle Position Estimation Based on Magnetic Markers: Enhanced Accuracy by Compensation of Time Delays. *Sensors* **2015**, *15*, 28807–28825. [CrossRef]
97. Encyclopedia Britannica Autobahn. 2018. Available online: https://www.britannica.com/technology/Autobahn-German-highway (accessed on 19 June 2023).
98. Thammanoon, P. Shenzhen, China, 16 August 2019: The Urban Trafic at Fuhua-Jintian Intersection at Shenzhen, China. Available online: https://www.shutterstock.com/image-photo/shenzhenchinaaugust-162019-urban-trafic-fuhuajintian-intersection-1481550272 (accessed on 19 June 2023).
99. Du, R.; Chen, S.; Li, Y.; Ha, P.Y.J.; Dong, J.; Anastasopoulos, P.C.; Labi, S. A Cooperative Crash Avoidance Framework for Autonomous Vehicle under Collision-Imminent Situations in Mixed Traffic Stream. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021; IEEE: Indianapolis, IN, USA, 2021; pp. 1997–2002.
100. Glaser, S.; Vanholme, B.; Mammar, S.; Gruyer, D.; Nouveliere, L. Maneuver-Based Trajectory Planning for Highly Autonomous Vehicles on Real Road with Traffic and Driver Interaction. *IEEE Trans. Intell. Transp. Syst.* **2010**, *11*, 589–606. [CrossRef]
101. Hu, L.; Zhong, Y.; Hao, W.; Moghimi, B.; Huang, J.; Zhang, X.; Du, R. Optimal Route Algorithm Considering Traffic Light and Energy Consumption. *IEEE Access* **2018**, *6*, 59695–59704. [CrossRef]
102. Jed, B. Two Roads: MS ChE or MBA? 2019. Available online: https://www.aiche.org/chenected/2019/05/two-roads-ms-che-or-mba (accessed on 19 June 2023).
103. Kummerle, R.; Ruhnke, M.; Steder, B.; Stachniss, C.; Burgard, W. A navigation system for robots operating in crowded urban environments. In Proceedings of the 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013; IEEE: Karlsruhe, Germany, 2013; pp. 3225–3232.
104. Aaron, M.; Edmund, B. With Chinese Help, Ethiopia Seeks to be Africa's Top Car Manufacturer. 2016. Available online: https://europe.autonews.com/article/20160604/ANE/160609945/with-chinese-help-ethiopia-seeks-to-be-africa-s-top-car-manufacturer (accessed on 19 June 2023).

105. Bai, H.; Cai, S.; Ye, N.; Hsu, D.; Lee, W.S. Intention-aware online POMDP planning for autonomous driving in a crowd. In Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26–30 May 2015; IEEE: Seattle, WA, USA, 2015; pp. 454–460.
106. Jiang, L.; Molnár, T.G.; Orosz, G. On the deployment of V2X roadside units for traffic prediction. *Transp. Res. Part C Emerg. Technol.* **2021**, *129*, 103238. [CrossRef]
107. Khayatian, M.; Mehrabian, M.; Andert, E.; Dedinsky, R.; Choudhary, S.; Lou, Y.; Shirvastava, A. A survey on intersection management of connected autonomous vehicles. *ACM Trans. Cyber-Phys. Syst.* **2020**, *4*, 1–27. [CrossRef]
108. Janai, J.; Güney, F.; Behl, A.; Geiger, A. Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art. *Found. Trends®Comput. Graph. Vis.* **2020**, *12*, 1–308. [CrossRef]
109. Funke, J.; Greiff, S. Dynamic Problem Solving: Multiple-Item Testing Based on Minimally Complex Systems. In *Competence Assessment in Education*; Leutner, D., Fleischer, J., Grünkorn, J., Klieme, E., Eds.; Methodology of Educational Measurement and Assessment; Springer International Publishing: Cham, Switzerland, 2017; pp. 427–443. ISBN 978-3-319-50028-7.
110. Chen, S.; Saeed, T.U.; Alqadhi, S.D.; Labi, S. Safety impacts of pavement surface roughness at two-lane and multi-lane highways: Accounting for heterogeneity and seemingly unrelated correlation across crash severities. *Transp. Transp. Sci.* **2019**, *15*, 18–33. [CrossRef]
111. UK Autodrive Pioneering Project Releases Final Report. 2019. Available online: http://www.ukautodrive.com/final-report/ (accessed on 19 June 2023).
112. Ye, L.; Yamamoto, T. Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput. *Phys. Stat. Mech. Appl.* **2018**, *512*, 588–597. [CrossRef]

sustainability

Article

Activity Duration under the COVID-19 Pandemic: A Comparative Analysis among Different Urbanized Areas Using a Hazard-Based Duration Model

Chunguang Liu [1], Xinyu Zuo [1], Xiaoning Gu [2], Mengru Shao [3,*] and Chao Chen [2,*]

1 Don School, International Education College, Shandong Jiaotong University, Jinan 250357, China; 213008@sdjtu.edu.cn (C.L.)
2 State Key Laboratory of Structural Analysis for Industrial Equipment, School of Automotive Engineering, Dalian University of Technology, Dalian 116024, China; guxiaoning@mail.dlut.edu.cn
3 Urban Planning and Transportation, Department of the Built Environment, Eindhoven University of Technology, 5600MB Eindhoven, The Netherlands
* Correspondence: m.shao@tue.nl (M.S.); chenchaocc@mail.dlut.edu.cn (C.C.)

Abstract: There have been significant changes in daily activities and corresponding durations since the outbreak of COVID-19. This study examines how the built environment factors and individual/household characteristics affect activity durations (e.g., shopping, social-related, hiking, and working) under the COVID-19 pandemic and analyzes the heterogeneity between different urbanized areas using the data of a Dutch national travel survey in 2020. A hazard-based duration model (e.g., the Cox proportional hazard model) was used to predict activity durations. Estimation results showed that the activity durations for different social groups varied under different geographical and policy conditions. In particular, women and seniors are more susceptible to the unprecedented pandemic, manifested in significantly shorter durations for work and hiking activities. In addition, couples with one or more children need to shorten their working hours and give more attention to their children due to the closure of nurseries and schools. Furthermore, the influences of built environment factors also present significant differences. A higher number of service facilities does not significantly foster the extension of hiking activity duration; however, this is the opposite among regions with more open green areas. Compared with previous studies on analyzing the influencing factors of activity durations, this study incorporated some unique variables (e.g., COVID-19 countermeasures and urban class) to consider the temporal and spatial heterogeneity under the particular pandemic period.

Keywords: built environment; COVID-19 countermeasures; activity duration; spatial heterogeneity; hazard-based duration model

Citation: Liu, C.; Zuo, X.; Gu, X.; Shao, M.; Chen, C. Activity Duration under the COVID-19 Pandemic: A Comparative Analysis among Different Urbanized Areas Using a Hazard-Based Duration Model. *Sustainability* **2023**, *15*, 9537. https://doi.org/10.3390/su15129537

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 5 May 2023
Revised: 1 June 2023
Accepted: 13 June 2023
Published: 14 June 2023

1. Introduction

Since the first known outbreak of COVID-19 in Wuhan, China, in December 2019, this novel coronavirus has spread rapidly around the world. Precisely because of the unexpected epidemic, many countries have suffered varying losses. By early October 2021, more than 233.2 million cases were confirmed in 215 countries, with more than 4.7 million deaths [1]. In order to limit the further spread of COVID-19 and reduce the number of deaths, most countries have imposed certain restrictions on their residents. Like other European countries, the Netherlands has also introduced some epidemic prevention and control measures in the last two years. These measures are expected to present an effective method to prevent and slow down the pandemic in the Netherlands. However, the practical effects depend on how society follows these rules. Subject to these kinds of stricter control measures, Shelat et al. [2] found that residents could be roughly divided into two groups: “COVID Conscious” and “Infection Indifferent”. In addition, the residents’ various activities could also provide valuable references for the government to adapt and update COVID-19 prevention measures further.

The COVID-19 pandemic and its countermeasures have disrupted daily life in many aspects. One of those apparent influences is on travel behavior and transportation mode choices. For instance, based on a worldwide web-based survey finding, Abdullah et al. [3] argued that during the COVID-19 pandemic, the primary trip purpose became supermarket shopping. In addition, private vehicles and active transport have become the first choice. Existing studies have largely focused on the analysis of the impact of pandemic preparedness policies on transportation and travel during the pandemic [3–6]. Moreover, there have been numerous studies examining the changes in activity duration during the COVID-19 pandemic [7–10]. However, the specific factors influencing activity duration in the context of the pandemic and whether the effects of built environment and socio-demographic factors on activity duration have been altered remain relatively unexplored. Therefore, in response to this gap, the paper aims to investigate the duration of various activities based upon proposed governmental control measures for COVID-19 and the built environment, primarily focusing on the Dutch context.

Due to the unique situation of the COVID-19 pandemic, the current paper investigates the activity duration by considering two aspects of drivers: the observed heterogeneity in socio-demographic characteristics and the built environment under COVID-19 countermeasures. The former has universal and observable variability, such as individual/household characteristics (e.g., age, income). The latter has its own unique features because the built environment factors (e.g., some public service facilities) under the pandemic can only be used with restrictions [11,12].

Firstly, many governments have adopted social distancing and family isolation measures to minimize the damage to health caused by COVID-19 [13]. Although measures are considered to reduce the transmission rate of this coronavirus, they can also endanger people's mental health, such as increasing the incidence of anxiety and depressive symptoms [14]. Therefore, influenced by this anxiety, people could be more inclined to participate in activities of a relaxing nature in their lives, such as travel and exercise [15]. However, not everyone is worried about COVID-19. Residents have differences in awareness of COVID-19 infection due to different socio-demographic characteristics, such as gender, education, and age [2]. Therefore, the latest research appears to show that individuals with different characteristics tend to have different activity features. Against this background, our research focuses on whether heterogeneous socio-demographic characteristics affect the activity duration under the COVID-19 pandemic. This exploration facilitates the identification of unique characteristics linked to different population groups, thereby equipping policymakers with new perspectives to develop customized and effective policies targeted at specific segments of the population.

Secondly, researchers have confirmed that exercise activities, such as sports and cycling, strongly correlate with the surrounding built environment for residents of all ages [16]. In addition, some studies have also confirmed that the built environment is inseparable from the duration of shopping activities [17]. As discussed in Koohsari et al. [18], the walkability of surrounding buildings will worsen with the improvement of the built environment, and the increase in population density will, in contrast, promote the sedentary behavior of residents. These factors have a potential impact on the duration of various types of activities. Furthermore, store opening hours have also changed due to the implementation of COVID-19 policies [19]. Consequently, the built environment and government policies have a profound joint influence on activity duration. Therefore, it is also necessary to analyze the relationship between the built environment, COVID-19 countermeasures, and the duration of various activities in life.

This paper investigates the impact of COVID-19 countermeasures taken by the Dutch government and the surrounding built environment of the community on activity duration based on a hazard-based duration model. The main contribution of this study does not lie in its methodology, as hazard-based duration models have already been successfully employed in studying event durations across multiple disciplines, such as biomedicine [20] and traffic safety analysis [21]. In relation to the topic of our study, hazard-based duration models have been deployed to understand household evacuation time behavior in the United States [22], as well as to investigate travel patterns concerning the use of new energy modes [23]. Additionally, related techniques have been applied to explore factors influencing social activity durations [24], gender differences in commuting activities [25], and walking durations during daily travel [26]. What sets our study apart is its pioneering approach, as it is the first, to the best of our knowledge, to employ this tool in analyzing the impact of various built environment factors and individual/household characteristics on activity durations during the COVID-19 crisis.

While this study focuses specifically on the Netherlands, and we acknowledge the substantial variations among countries in terms of their culture, urbanization, government-implemented COVID-19 countermeasures, and daily travel patterns [5,6,27], we nonetheless also note that many other countries have experienced similar changes in behaviors during the pandemic, akin to those observed in the Netherlands [28]. Therefore, considering the shared features of reduced travel demand and ongoing debates on COVID-19 countermeasures, we expect that that our findings will provide valuable insights extending beyond the specific context of the Netherlands.

The rest of this paper is organized as follows. Section 2 reviews the existing relevant literature on behavior changes under the COVID-19 pandemic and methodologies. Section 3 explains the data, study area, and research time. Section 4 develops the hazard-based duration model to derive hazard function distributions for various activities. Estimation results for the considered factors and significant findings are then elaborated in Section 5. Finally, the paper gives conclusions and recommendations for future research.

2. Literature Review

In the past two years, extensive and continuous measures have been announced by governments worldwide to stem the spread of COVID-19 [5,29]. Nevertheless, a problem is that people's lives have been strongly shaped, and these effects are unprecedented. Significant influences of COVID-19 countermeasures and built environment factors on daily activities have been found in prior studies, presenting the differences among different communities and social groups [3,4,11,12,30,31]. Since the objective of this paper is to examine the relationship between socio-demographic, built environment factors, and activity durations under the COVID-19 pandemic and to explore the spatial heterogeneity of influences across different urbanized areas, the review mainly focuses on the pandemic and the resulting activity duration features. The related aspects could be summarized as follows: (1) the impact of COVID-19 on people's daily life; (2) activity duration and its influencing factors; (3) methodology; and (4) research gaps.

2.1. Life Changes under the COVID-19 Pandemic

Amid the continual spread of COVID-19, Thunström et al. [29] and Cohen and Kupferschmidt [32] have noted the necessity and feasibility of the requirement for a part of the community to stay at home such that the social distance can be maintained and the strategy of "flattening the transmission curve" can be ensured. Indeed, such large-scale travel restrictions have had significant negative effects on individual travel and economic activities of enterprises. In particular, the unprecedented pandemic has severely disrupted travel patterns in almost all countries worldwide. Taking the United States as an example, the decline in accommodation orders in every American state provides evidence for the shrinking of traveling [30]. Studies have also shown that, due to the fear of COVID-19, even a low risk of COVID-19 infection can significantly affect the turnover rates in an area. In addition, the decline in the turnover rates was also influenced by the citizens' desire to follow governmental regulations [33]. Furthermore, due to the closure of recreational, educational, and workplace facilities and the local or international travel restrictions, individuals were forced to spend more time in their habitations. Consequently, they began to find ways to replace or supplement their non-family activities with family activities [4].

Some studies also indicated that, from the long-term point of view, individual voluntary behaviors rather than government regulations might be more critical in controlling COVID-19 transmission in Western countries [34]. Specifically, compared with citizens in other European countries who were forced to stay at home, the Dutch government relied a great deal on amoral appeals, encouraging citizens to stay at home and maintain a 1.5 m social distance as much as possible. All public spaces, from parks to beaches, were kept open in the early days of the outbreak of COVID-19 [35]. However, as the number of infections increased, the Netherlands also adopted "intelligent lockdown". During that period, cafés, restaurants, gyms, and schools were forced to close [36]. Obviously, young people were free to join clubs and watch movies, women were free to participate in shopping activities, and men were free to practice or watch football before the pandemic [37]. Under the lockdown policies, people's activity duration had to be adjusted passively (i.e., from March to July 2020).

Meanwhile, an extensive body of research has been dedicated to investigating the impacts of the COVID-19 pandemic and its associated policies on travel behavior, placing specific emphasis on scrutinizing the spatial and social heterogeneity. Utilizing extensive cellphone big data collected from regular metro users, Liu and Zhang [5] conducted a comprehensive study to examine the socially and spatially heterogeneous impacts of mobility intervention policies on daily metro transit use in Shenzhen, China. Concurrently, Zhang and Li [6] adopted a similar methodology to investigate the causal effects of mobility intervention policies on urban park visits during the COVID-19 pandemic. Furthermore, employing spatial regression models, Li et al. [38] explored the utilization of urban parks before and during the pandemic in Guangzhou, China, and identified spatially uneven effects, underscoring the importance of considering sub-city-level differences (i.e., spatial heterogeneity) in city planning. These aforementioned studies collectively demonstrate the presence of social and spatial heterogeneity in influencing various factors.

2.2. Activity Duration

Previous studies generally divided activities into work-related and non-work-related activities and indicated that the time allocation of work-related activities has better predictability and lower dependence on personal daily spatiotemporal constraints [39]. Compared to work-related activities, the modeling of non-work-related activities is more complicated, and these complexities could be attributed to the flexibility, variability, and randomness of such behaviors [40]. In addition, the examination of relevant literature showed that the non-work-related activities have great diversity, making diversity seeking a critical factor in the choice of non-work-related activities (e.g., shopping activity, travel activity, social activity). The duration of these activities has been extensively investigated in the literature [24].

Bláfoss et al. [41] found that the durations of work-related activities and non-work-related activities were negatively correlated and varied with the changes in age and physical strength. Analogously, Sreela et al. [42] analyzed the duration of shopping activities based on parametric and semi-parametric hazard-based models. In their research, the impacts of household status, gender, age, travel cost, and activity start time were considered. In terms of economic conditions, Sarangi and Manoj [43] developed a multivariate Probit model to evaluate the decision making for activity participation and analyzed the duration of various activities. Furthermore, with time and budget constraints, Dane et al. [44] proposed an activity-based model to simultaneously predict the duration of several outdoor leisure activities.

The effects of built environment factors on activity duration have also been examined in previous studies [31]. Some researchers found that the built environment factors are closely related to outdoor leisure activities [17,23,45,46]. In particular, various studies have highlighted the effects of the built environment on the duration of outdoor activities and sedentary behavior of young and old adults [47,48]. Based on spatial hazard modeling, Anastasopoulos et al. [23] examined the impacts of distance of different facilities on the duration of hiking activities. Clark et al. [45] also found that the changes in outing activities are significantly affected by life events, spatial context, and individual perceptions (e.g., attitudes towards the built environment). Hahm et al. [17] examined the causal relationship between the built environment, walking duration, and activity duration of shopping activities using GPS data in Seoul retail districts. Based on the national travel survey data of the Netherlands, Wang et al. [46] examined the impacts of the built environment and natural environment factors on the duration of recreational activities. In the context of the pandemic, Park et al. [31] conducted a study examining the impacts of COVID-19 and built environments on daily community living activities. Their research specifically concentrated on individuals with disabilities, who are particularly susceptible to the effects of the pandemic.

2.3. Methodologies

In recent years, the growing focus on studying time allocation has sparked a quest for flexible and tractable modes, exemplified by the emergence of the multiple discrete continuous extreme value (MDCEV) model as a prominent approach [49]. Originating from the foundational concept of maximizing individual utility, multiple discrete continuous (MDC) structures have evolved into a framework to model decisions related to activity participation and time allocation while accounting for budget constraints. These methodological advancements have yielded noteworthy contributions in understanding the dynamics of activity participation [50]. More recently, the MDC model, which emphasizes the heterogeneity in socioeconomic characteristics and individual activity duration, has been further developed into a choice model with latent variables [51]. Although some MDC variant models (e.g., multiple discrete continuous extreme value models) improved the variable handling method, in most cases, the MDC models still treat the duration of activity as a discrete variable [52]. In addition, it is important to highlight that, owing to the data prerequisites involving both discrete choice and continuous consumption decisions, MDC models are commonly applied with stated preference data as opposed to revealed preference data [53]. Conversely, the data used in this study fall under the category of revealed preference data, thereby posing challenges in utilizing the MDC model to examine the determinants influencing activity duration.

Due to the dynamic characteristics, non-work-related activities have inherent randomness, leading to uncertainty in estimating activity duration. The fuzzy set theory has been proved to be an effective method to deal with the uncertainty in construction projects [54]. Salah and Moselhi [55] modeled emergency activities in projects based on the fuzzy set theory and used the digital fuzzy number to express the total duration of each activity. However, the fuzzy set theory cannot assign proper weights to the estimation. In addition, the application of the average method may also cause incorrect membership functions, a core element of fuzzy set theory [56]. Consequently, Gładysz [57] proposed a hybrid probability distribution model that employed beta distribution and generation probability distribution of fuzzy variables to model task durations. In their model, the deterministic level was also incorporated into the most likely time estimation. However, this particular model is better suited for modeling activities within construction projects rather than individual travel activities. Therefore, it is crucial to find a method that can effectively model and investigate the factors influencing the variations in daily activity durations.

In addition to the aforementioned methods, other commonly used approaches in the literature to investigate the correlation between the built environment and physical activity duration include multilevel Tobit regression models [58], structural equations [16], multiple regression models [31], and spatial regression models [38]. However, there is an important concept that needs to be highlighted, namely, the dependence of the probability of an activity's end on the duration of the activity. In other words, the likelihood of ending an activity is influenced by the amount of time already spent on the activity. However, these models do not take this characteristic into account. Therefore, it is crucial to find a model that can incorporate this concept in studying duration-based data.

The use of hazard-based duration models to study the time of occurrence for an event has been a subject of research. Due to the flexibility of accounting for censored observations of duration, this kind of model has been widely used in the analysis of duration data. In particular, the proportional hazard model is the most commonly used method in duration research [59]. For instance, using the proportional hazard model, Yee and Niemeier [60] analyzed the duration data obtained from Puget Sound Transportation Panel surveys. Additionally, based on a hazard-based econometric model, Anastasopoulos et al. [61] analyzed the effect of influential factors on trip distance. Pang and Krathaus et al. [21] adopted a hazard-based duration model considering multiple layers of heterogeneity to analyze accident occurrences during snow events. It needs to be emphasized that heterogeneity is a critical factor in duration modeling. The most commonly used method in the literature to consider this factor is to incorporate an unobserved random term. Another alternative that could include heterogeneity is to separately estimate parameters for different segments of individuals using the underlying factors. Relative to other econometric models (e.g., standard regression), the Cox proportional hazards model could reasonably consider the effect of heterogeneity. Therefore, in this paper, the Cox proportional hazards model is adopted to examine the relationship between activity duration, socio-demographic characteristics, built environment factors, and some representative COVID-19 countermeasures under the COVID-19 pandemic.

2.4. Research Gap

The review of existing works presented above clearly shows missing knowledge of activity duration under an emergency. Most studies only consider the relationship between socio-demographic characteristics and activity duration, ignoring the impact of environmental factors such as the built environment [26]. Considering the differences in the built environment due to COVID-19 pandemic countermeasures, further understanding the heterogeneous impacts of built environment factors and COVID-19 countermeasures on activity duration is needed. Furthermore, studies on the relationship between activity duration and built environment factors only focus on a single area and fail to explain the differences in activity duration between different urbanized areas. The spatial heterogeneity remains underexplored.

Therefore, to fill these research gaps, this study explores the duration patterns of various activities by analyzing the Dutch national travel survey and physical neighborhood data in 2020. Rather than defining complex land-use mix indexes, like in the literature, we directly used lower-level land-use variables to measure the built environment in this study. In addition, the impacts of COVID-19 countermeasures are also considered for the first time. Moreover, we also examined the interaction effects between built environment factors and COVID-19 countermeasures, e.g., the number of opened stores under the COVID-19 pandemic. Finally, the spatial heterogeneity in the duration of various activities for Dutch society was also recognized using the hazard-based duration model. Results obtained in this paper could serve as a reference for policymakers to understand the effectiveness of COVID-19 countermeasures in different areas and take timely steps to better control the spread of COVID-19.

3. Data and Study Area

3.1. Study Area, Time Span, and Dependent Variable

This paper considers all of the Netherlands as the study area. As different urban contexts are found in different Dutch cities, the individual travel behaviors are bound to show different traits. In the literature, the regional division of the Netherlands is usually made based on municipalities, towns, or districts. Statistics Netherlands also takes these classifications to publish statistics [36,46]. However, as an important identification code for location, the classification according to postcode areas has also become popular in recent years. Therefore, this paper aggregates all data at the postcode level for the analysis.

The data employed in this study were retrieved from different data sources. In particular, being the most critical dependent variables, the activity duration was collected as part of the Dutch national travel survey conducted from January to December 2020 [62]. The survey includes data from 62,940 respondents, about 0.38% of the population aged six years and older. Respondents were asked to keep a record of where they visited, for what purpose, with what means of transport, and how long the activity lasted for one particular day of the year. Additionally, information of individual/household characteristics, such as age, gender, and driving license, was also collected. The location data were registered at the 4-digit postcode level, which was used to link with the data at the neighborhood level. The data contained a total of 4072 4-digit postcode areas (see Figure 1).

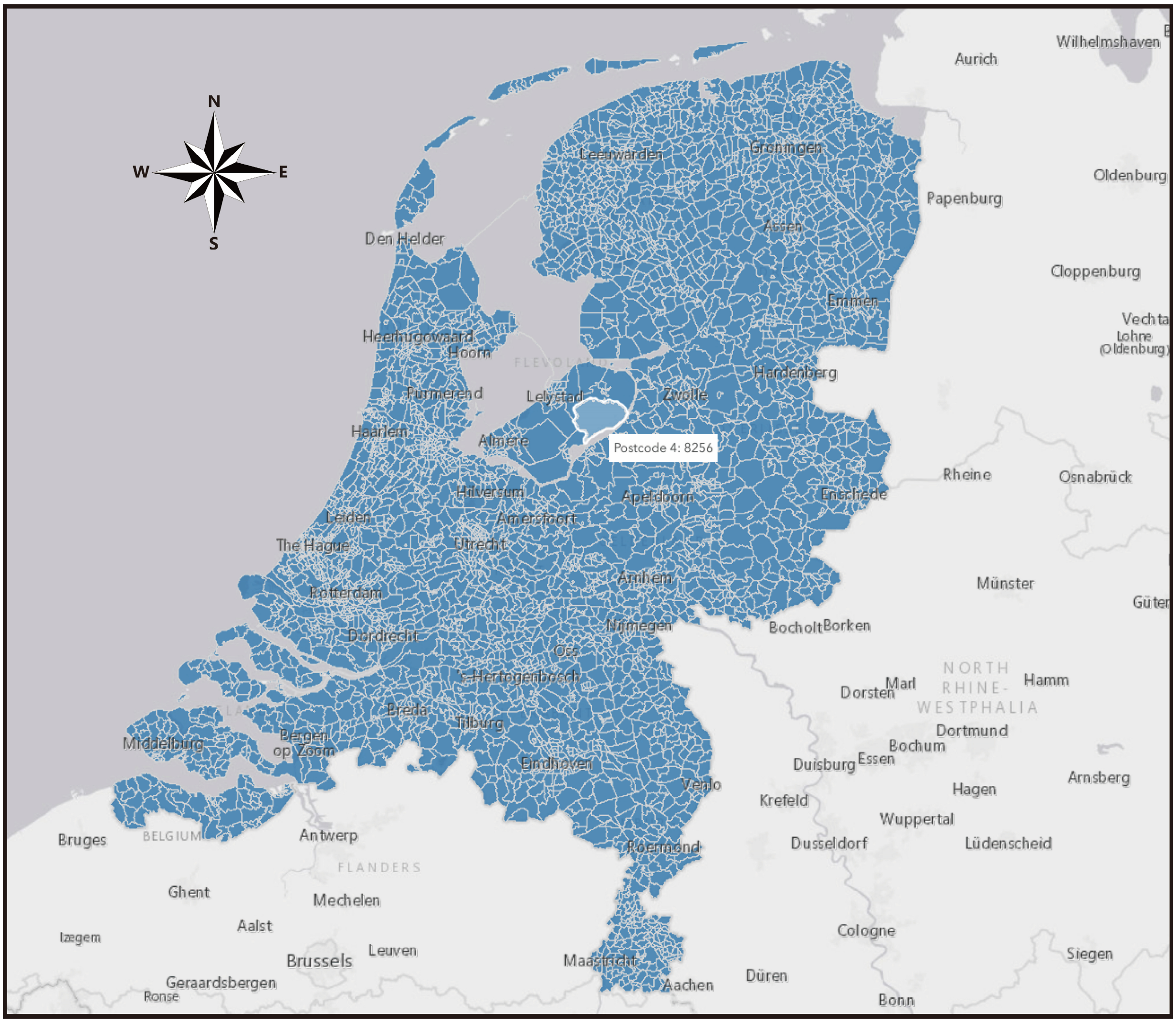


Figure 1. Study area of the Netherlands (four-digit postcode areas).

A detailed classification of activities (e.g., nine kinds of activities) was used in the survey, and it is available from the corresponding author upon request. However, considering the potential influence of COVID-19 policies on the possibility of various activities, these activities were reclassified into four major activity classes: daily shopping, social-related, hiking, and in-office work. Furthermore, to ensure that all activities examined originated from the home location, activities departing from non-home addresses were excluded from the analysis. The resulting sample size available for the present investigation included

12,392 shopping activities, 6481 social-related activities, 8870 work-related activities, and 4142 hiking activities. Based on the collected data, the average duration of activities was found to be 43.7 min for shopping activity, 129.7 min for social-related activity, 315.7 min for work-related activity, and 167.9 min for hiking activity.

3.2. Explanatory Variables

The selection of the influential factors was largely based on the previous work on activity analysis [26]. In Tables 1 and 2, we provide a descriptive summary of the variables. Nine variables related to individual/household characteristics were included to capture the effects of social demographics on activity duration. In addition, several groups of neighborhood characteristics were also considered to explore the impacts of built environment factors on activity duration. Furthermore, four types of COVID-19 countermeasures were included to assess the effectiveness of various Dutch governmental measures on reducing regular activity duration.

3.2.1. Social Demographical and Neighborhood Characteristics

Among individual/household characteristics, the effects of gender, age, driver's license, education level, migration background, income status, household status, the number of vehicles in the household, and the number of young members in a family were examined. Some variables have continuous values while others are categorical. For categorical variables, the first category was considered as a reference.

The neighborhood data were extracted from the Dutch Central Bureau of Statistics [63–66]. Five aspects were considered, including density, facilities, green space, land-use mix indexes, and an urban class variable. The density variables include population density and neighborhood address density. Facilities contain a range of facility–distance variables and the number of facilities available within a 5 km radius. Note that the facility–distance variables represent the average distance from the neighborhood center to its nearest facility. In this study, various types of facilities are considered, namely health and well-being facilities (e.g., hospitals), retail facilities (e.g., department stores), catering facilities (e.g., restaurants), education facilities (e.g., primary schools), traffic and transport facilities (e.g., train stations), and leisure and cultural facilities (e.g., museums and music venues). For a detailed explanation of these facility variables, please refer to Tables 1 and 2. In addition, as discussed earlier, people under the pandemic may become more willing to use open and clear space to avoid crowds and reduce the infection risks. Thus, a wide variety of green space variables, including open green spaces (e.g., green park space, forest, and open nature space) and daily recreational spaces (e.g., sport space and recreation area), were included.

Moreover, we tested a wide range of land-use mix indexes [67], including the Balance Index, the Entropy Index, and the Herfindahl–Hirschman Index. However, none of these indicators showed significant improvements in the model goodness of fit and interpretability. Thus, we used the percentages of specific land-use types within an area (e.g., land use for traffic, land use for residential buildings, land use for agriculture, and land use for recreation). Finally, an urban class variable was used to represent the level of urbanization within an area. A detailed description of these built environment factors is provided in Tables 1 and 2.

It is also important to point out that the neighborhood data do not entirely match the travel survey data presented in Section 3.1 because the travel data were identified based on the 4-digit postcode, while the neighborhood data were derived from the Esri-open postcode plane with 8 digits. Indeed, the 4-digit postcode area represents a higher level of spatial aggregation than the 8-digit neighborhood postcode. For further analysis, the neighborhood data were aggregated to the 4-digit postcode level. Based on the final candidate variable dataset, some statistical indicators for the candidate variables are also presented in Table 2.

Table 1. The list of variables.

Variables	Explanation of Variables
Dependent variable	
Shopping activity duration	The duration of shopping activities for a specific activity and person
Social-related activity duration	The duration of social-related activities for a specific activity and person
Hiking activity duration	The duration of hiking activities for a specific activity and person
Work-related activity duration	The duration of work-related activities for a specific activity and person
Independent variable	
Socio-demographic characteristics	
Gender	The gender of the person
Age	The age class of the person
Driver's license	Whether the person has a valid driver's license
Education	The education level of the person
Migration background	The migration background of the person
Household income level	Deviation from the Dutch low-income threshold
Household status	The household composition/type of the person (including single household, couple, couple +child(ren), couple +child(ren)+other(s), couple +other(s), single-parent household +child(ren), single-parent household +child(ren)+other(s), and another household
Car number	The number of cars in the household
Young members in a family	Number of household members younger than 6 years
Built environment: density variables	
Population density	Population per square kilometer
Neighborhood address density	The average number of addresses per square kilometer
Built environment: facility variables	
Facility distance	The average distance from the center of the neighborhood to the near-daily or non-daily facilities, such as supermarket, daily goods stores, etc.
The number of facilities	The number of daily/non-daily facilities within 1/3/5 km distance by road for all residents of an area
Built environment: green space variables	
Open green space	The percentage of the area of each type of open green space (including green park space, forest, open nature space, and inland water area)
Recreational space	The percentage of the area of each type of recreational space (including sports space and daily recreational space)
Urban class variable	
Urban class	The urbanization level defined by the Dutch Central Bureau of Statistics (including very strongly urban, highly urban, moderately urban, slightly urban, and not urban)
COVID-19 countermeasure variables	
Social distance rule	Whether the social distance rule is in effect
Work from home suggestion	Whether the suggestion of working from home if possible is in effect
Entertainment open rule	Whether restaurants, bars, and other recreational venues are allowed to open
Face mask requirement	Whether the face mask requirement is in effect

Table 2. The descriptive analysis of the final candidate variable dataset.

Variables	Min	Max	Mean	SD	Unit
Dependent variable					
The duration of shopping activity	5.00	180.00	43.70	40.81	min
The duration of social and recreational activity	5.00	1115.00	129.67	112.65	min
The duration of hiking and hiking activity	5.00	600.00	167.89	121.51	min
The duration of work activity	10.00	1095.00	315.74	208.38	min
Independent variable					
Built environment: density variables					
Population density	5.00	26,986.00	5017.33	4167.20	per km^2
Neighborhood address density (EAD)	20.00	11,545.00	2123.12	1933.73	per km^2

Table 2. *Cont.*

Variables	Min	Max	Mean	SD	Unit
Built environment: facility variables					
Supermarket distance (DTLS)	0.20	9.30	0.96	0.68	km
Daily goods store distance	0.04	7.75	0.84	0.62	km
Café distance (DTCAFE)	0.04	10.25	1.22	1.01	km
Cafeteria distance (DTCAFETERIA)	0.04	8.60	0.81	0.68	km
Restaurant distance	0.02	8.30	0.80	0.60	km
Hotel distance	0.09	13.30	2.19	1.75	km
Daycare distance	0.15	8.80	0.75	0.49	km
Out of school childcare distance	0.15	9.85	0.83	0.57	km
Primary education distance	0.20	9.00	0.85	0.48	km
Secondary education distance	0.20	19.44	2.15	2.00	km
Vmbo distance *	0.20	19.44	2.26	2.02	km
Havo vwo distance *	0.31	38.95	2.88	2.85	km
Train station distance (DTTS)	0.30	57.80	4.59	5.40	km
Transfer station distance (DTITS)	0.60	111.20	11.56	11.98	km
Main highway distance (DTMHE)	0.10	39.40	1.80	1.11	km
Cinema distance (DTCINEMA)	0.27	56.50	5.47	5.01	km
Library distance	0.30	17.40	1.87	1.30	km
Museum distance	0.29	24.25	3.52	2.79	km
Performing arts distance	0.20	36.10	4.16	3.95	km
Attraction distance	0.50	51.73	5.27	4.33	km
Number of supermarkets within 5 km	0.00	161.68	28.88	32.24	
Number of daily goods stores within 5 km (NOSTORES)	0.00	1015.97	138.52	206.20	
Number of cafés within 5 km	0.00	711.30	77.27	139.48	
Number of cafeterias within 5 km	0.00	998.06	126.90	196.42	
Number of restaurants within 5 km (NOREST)	0.00	1826.15	177.25	323.49	
Number of cinemas within 5 km (NOCINEMAS)	0.00	11.98	1.70	2.39	
Number of daycare facilities within 5 km (NOOFSC)	0.00	191.30	43.16	41.41	
Number of primary education facilities within 5 km	0.10	120.15	27.69	24.01	
Number of museums within 5 km	0.00	42.88	4.58	8.08	
Number of performing arts facilities within 5 km	0.00	39.35	3.94	6.91	
Number of attractions within 5 km	0.50	51.73	5.27	4.33	
Built environment: green space variables					
Business park	0.00	74.84	6.94	8.28	%
Park	0.00	41.54	4.86	5.38	%
Sports area	0.00	36.24	2.51	3.40	%
Allotment garden	0.00	68.63	0.42	2.22	%
Recreational area	0.00	46.66	0.33	1.67	%
Day recreation area	0.00	24.14	0.43	1.65	%
Forest and open natural terrain	0.00	76.26	4.32	8.85	%
Built environment: lower-level land-use variables					
Land use for traffic	0.07	66.14	16.28	15.38	%
Land use for retail and hospitality area	0.00	72.13	4.68	9.03	%
Land use for public facilities	0.00	19.23	0.92	2.15	%
Land use for socio-cultural provision	0.00	53.85	2.91	4.68	%
Land use for residential buildings	0.00	98.33	40.39	20.07	%
Land use for agriculture	0.00	95.99	17.78	22.25	%
Land use for recreation	0.00	76.47	8.56	7.68	%
Urban class variable					
Urban class	0.00	5.00	2.69	1.36	n/a
COVID-19 countermeasures					
Social distance rule	0.00	1.00	n/a	n/a	n/a
Work from home suggestion	0.00	1.00	n/a	n/a	n/a
Entertainment open rule	0.00	1.00	n/a	n/a	n/a
Face mask requirement	0.00	1.00	n/a	n/a	n/a

* Vmbo and Havo vwo represent secondary professional education in The Netherlands; n/a is the abbreviation of not applicable.

3.2.2. COVID-19 Countermeasures

As in many other countries, unprecedented packages of COVID-19 countermeasures, namely intelligent lockdown and hard lockdown, were also announced in the Netherlands. Figure 2 summarizes a timeline of COVID-19 countermeasures announced by the Dutch government over the whole of 2020. Besides this, the number of new cases daily (along with the policy packages) during the same period is also presented. Policies such as working from home, banning large-scale events, and social distancing undoubtedly have a significant impact on activity choice behavior and activity duration.

Therefore, this study included the variables presenting the information on the Dutch countermeasures to consider the effects of COVID-19 countermeasures on activity duration. In particular, several representative policies, e.g., suggestions for working from home, physical distancing requirement, mandatory mask wearing, and entertainment closure, were included as control variables of our models. The selection of the four countermeasures could also be attributed to the fact that these four measures are some of the most well-adopted response strategies worldwide, making the result comparison of different countries in future studies easier and more convenient [12,36,68].

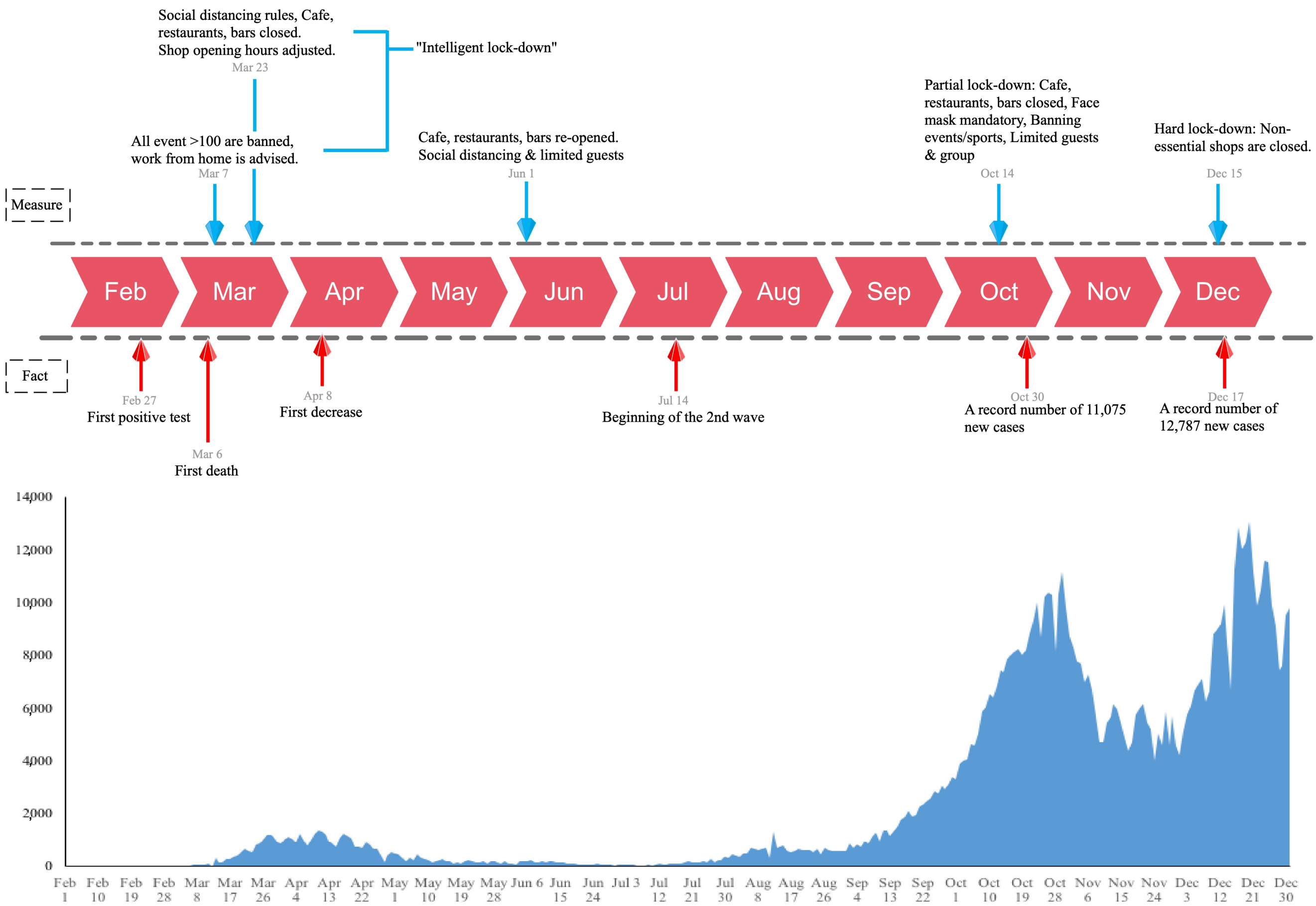


Figure 2. Measures against COVID-19 and the number of new cases daily during the same period.

3.3. Descriptive Statistics

In short, four dependent variables and six groups of explanatory variables are considered in this paper. Table 2 shows a detailed overview of the variable classification and the descriptive statistics of all involved variables.

4. Methodologies

In this study, the activity durations under the COVID-19 pandemic were modeled based on hazard-based duration models. Specifically, a semi-parametric model, known as the Cox proportional hazards model, was developed to examine the relationship of various covariates with the activity duration data. Using this model, the survival probabilities of the duration data (or the conditional probabilities of ending activities at a specific time t) could also be estimated and acquired [24].

4.1. Survival Analysis

In hazard-based duration models, a hazard function $h(t)$ gives the rate at which an event (e.g., ending of an activity) occurs during a specific time interval, e.g., t to $t + \delta$, given that the event has not occurred. This conditional probability of an activity's end is critical in the model because the probability that an activity terminates depends on the length of time the activity has lasted. For instance, when examining activity durations, the probability of an individual ending an activity is contingent upon the time the individual has already spent on that activity. Due to this feature, the hazard-based duration model was selected in this study, rather than other statistical models (e.g., partial least square regression), to examine the determinants of activity duration. In this study, it is assumed that the activity duration starts when an activity is about to be executed and ends when the activity is actually performed. The hazard function can be expressed mathematically in the following equation.

$$h(t) = \lim_{\delta \to 0^+} \frac{P(t \leq T < t + \delta | T \geq t)}{\delta} = \frac{f(t)}{1 - F(t)} \tag{1}$$

where t is a specific time and T is a non-negative continuous random time variable with a probability density function $f(t)$ and cumulative distribution function $F(t)$. Note that the function $F(t)$ defines the probability that an activity is completed before time t The mathematical expressions for the two functions are given by Equations (2) and (3).

$$F(t) = P(T < t) \tag{2}$$

$$f(t) = \frac{F(t)}{t} \tag{3}$$

In addition to the hazard function, the survival function is another important concept in hazard-based duration models [69]. This function is defined as the probability that the activity duration is greater than or equal to a specific time t (see Equation (3)).

$$S(t) = P(T \geq t) = 1 - P(T < t) = 1 - F(t) = \frac{f(t)}{S(t)} \tag{4}$$

The most frequently used method to estimate the survival function in the literature is the Kaplan–Meier method [70]. This non-parametric model exploits the product limit of conditional probabilities to estimate the approximate survival function. A general formula for the Kaplan–Meier estimator is shown below.

$$S(t_k) = \prod_{i=1}^{K} \frac{r(t_i) - d(t_i)}{r(t_i)} \tag{5}$$

where $r(t_i)$ represents the number of individuals at risk of ending their activity in time period i, $d(t_i)$ is the number of individuals terminating their activity in time period i, and K is the number of all time periods. However, it should be pointed out that the non-parametric models cannot involve any other covariates except for the duration variable.

4.2. Hazard-Based Duration Model

The effect of covariates on the activity duration can be accounted for by two alternative hazard models: the accelerated hazard (AH) model and the proportional hazard (PH) model. Unlike the PH model, which assumes the covariates have a constant effect on an unspecified baseline hazard function, the AH model requires a prior distribution (e.g., exponential, Weibull, log-logistic, etc.) for the hazard function. Previous studies also pointed out that the approach will lead to inconsistent estimates if the assumption of the distribution of hazard function is incorrect [71,72]. Alternatively, the semi-parametric hazard model is convenient when there is little or no knowledge on the hazard functional form available; in addition, the parametric assumption of the effects of covariates is still retained. For a detailed discussion of hazard-based duration models, see Hensher and Mannering [69] and Bhat [73]. In this paper, in order to address the above issues, the most frequently used approach in modeling duration data from biostatistics and economics, namely the Cox proportional hazards model, is employed. For this model, the hazard function is generally expressed as follows.

$$h(t, \boldsymbol{X}) = h_0(t) \exp\left(\boldsymbol{\beta}^T \boldsymbol{X}\right) \tag{6}$$

where $h_0(t)$ represents the baseline hazard function, i.e., a non-parametric model that does not need to be nominated as having a specific distribution. The other part is a parametric model, in which the exponential of a linear combination of covariates and corresponding parameters is involved. $\boldsymbol{X}$ is a vector of attributes describing the socio-demographics of individuals and the built environment. The vector $\boldsymbol{\beta}$ is a set of unknown parameters to be estimated. Note that this model could be estimated using standard maximum likelihood methods. Combining the estimated parameters and hazard ratios, this model could provide an adequate explanation for the duration data.

Based on the *t*-statistic, variables introduced in Table 2 were included in the final models if they were significant in any single activity duration model. In particular, a variable that had a significant effect on one activity was also included in other activities, even if this variable was indeed insignificant in that model. The main reason for this action was to fairly compare the effect of a variable on different activities (i.e., how the variable influences duration differently across different activities).

5. Results

5.1. Results of Survival Curve Analysis

As discussed earlier, this study aims to analyze the heterogeneity of activity duration between different urbanized areas. In this case, the distribution of activity duration is first examined. As an initial step, the survival curves for different urban classes were first estimated and depicted (see Figure 3). Collectively, this visual assessment provides an opportunity not only to examine the distributional characteristics of different activity durations but also to determine whether the differences in activity durations between different urbanized areas are sufficiently significant to warrant further investigation.

The survival curves presented in Figure 3 are for four different types of activities. As there are some overlaps for the survival functions of different urbanized areas in some specific activities, we also present a fitted hazard curve (see Figure 4) to jointly analyze the differences in activity durations between different urbanized areas. The findings deduced from these analyses are summarized as follows.

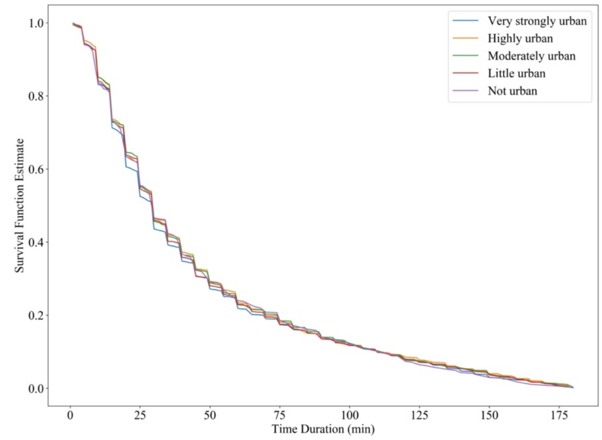
A general finding is that the activity behaviors (e.g., shopping, socializing, hiking, and work) are similar across Dutch citizens in different urbanized areas because the trends of all graphs for each activity are almost identical. However, as can be seen, the significance of the drops in the value of the survival function for different urbanized areas is higher among social, hiking, and working activity than shopping activity.

More specifically, Figure 3a indicates that the drop in shopping activity duration in highly urbanized areas is more remarkable than other areas in periods under 75 min, but the opposite in periods over 125 min. In that time period, the shopping activity durations for citizens living in non-urbanized areas fell at the fastest rate. One potential explanation for this finding is that the higher density of shopping facilities in highly urbanized areas leads to residents having relatively shorter shopping durations, mainly concentrated within 75 min. In contrast, residents in non-urbanized areas, confronted with the inconvenience of shopping, engage in longer shopping durations to meet their needs for an extended period. Consequently, they tend to make bulk purchases to fulfill their future requirements, resulting in a relatively longer shopping duration. In addition, as shown in Figure 3a, the probability of continuing the shopping activity declines more steeply in the periods under 50 min than other periods for all urbanized areas. The probability of continuing the shopping activity after 50 min was only about 23%, and there is 50% likelihood that Dutch citizens end their shopping activities in 29 min. This value is slightly lower than the average time Europeans spend on shopping activities, but the finding is consistent with our experience (https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20181123-1, accessed on 4 May 2023).

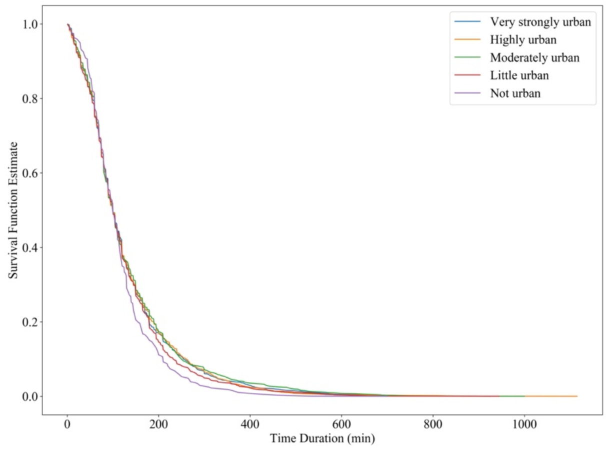
Similarly, the survival functions for social-related activities are summarized in Figure 3b. It can be seen that the main drops in the survival functions are seen at up to 4 h, and the rate of reduction rapidly decreases after this period. In particular, the probability of social-related activities longer than 4 h is 11.3%, 11.7%, 11.3%, 8.7%, and 6.2% for Dutch citizens in very strongly urban, highly urban, moderately urban, slightly urban, and non-urban areas, respectively. We can also observe that the fall in the survival function values is more remarkable in less urbanized areas (e.g., slightly and not urban) than in other areas. One plausible explanation for this finding is that residents living in less urbanized areas, where social facilities are limited in availability, need to travel a greater distance to engage in social activities. The relatively longer return journey, coupled with the increased risk of infection during such social interactions, prompts them to prefer concluding their activities at an earlier time.

Regarding the survival functions for hiking activities, Figure 3c indicates that the probability of continuing hiking activity decreases almost linearly in periods under 300 min. In addition, it is interesting to note that the decline in hiking activity duration in non-urban areas has a relatively high magnitude. This result implies that a wide-open area does not lead to a significant increase in hiking activities. Quite the opposite, the hiking activity duration is much higher in slightly urban areas, especially when the periods are over 200 min. Specifically, compared to the average hiking time provided in Table 2, the probability of continuing hiking activity for longer than that time is 47.4% and 38.7% for citizens in slightly urban areas and non-urban areas, respectively. This finding raised a question regarding the shared experience that more open green areas result in more hiking activities. It seems possible that the lack of essential public facilities (e.g., shops and restaurants) might be the cause of the opposite results in non-urban areas.

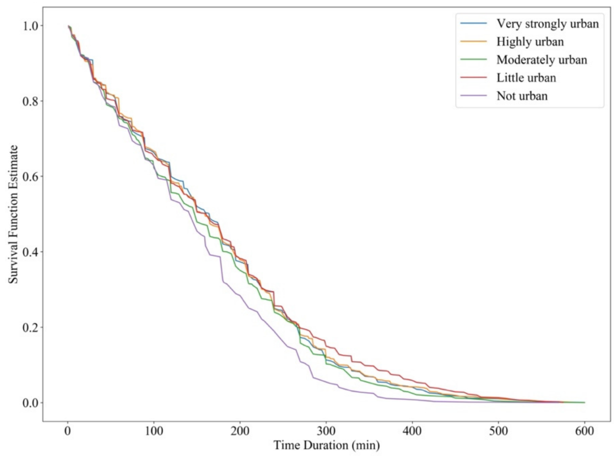
Particular attention should be given to the survival functions of working activities, which have a relatively interesting distribution pattern. Unlike the distribution pattern of other activities, the S-shaped survival function for work activities shows that the reduction rate is relatively high in periods under 2.5 h and between periods of 7.5 h and 12 h. According to Figure 3d, the work activity duration for nearly 40% of the respondents lies between 2.5 h and 12 h, and the probability of work activities longer than 600 min (i.e., 12 h) per day is negligible. These values correspond well with findings from most previous studies (https://www.spica.com/blog/working-regulations-netherlands, accessed on 4 May 2023). For example, according to the latest OECD Better Life Index, most full-time employees in The Netherlands work 36–40 h per week (e.g., 430 min to 480 min per day). In addition, Dutch part-time work limits the working time to between 12 and 36 h per week. The data also provide indirect insights into the underlying reason for the S-shaped pattern observed in the survival function for work activities. The implementation of pandemic-related policies has led to a decrease in the operating hours of numerous businesses or service establishments and, in some cases, complete closures. Consequently, there has been a reduced demand for short-term part-time positions, and there is no need for full-time employees to extend their working hours. Instead, they are simply expected to fulfill their prescribed work hours, contributing to the relatively gradual decline observed in the middle part of the survival function for work activities. Furthermore, it can also be observed that the reduction rate of the survival function decreases with the increasing level of urbanization, indicating that citizens in highly urbanized areas tend to work longer than those who live in non-urban areas. This finding is as expected because the increasingly competitive environment in urban areas puts many people under great pressure and, consequently, they need to work hard for a better quality of life [74,75].



(**a**)

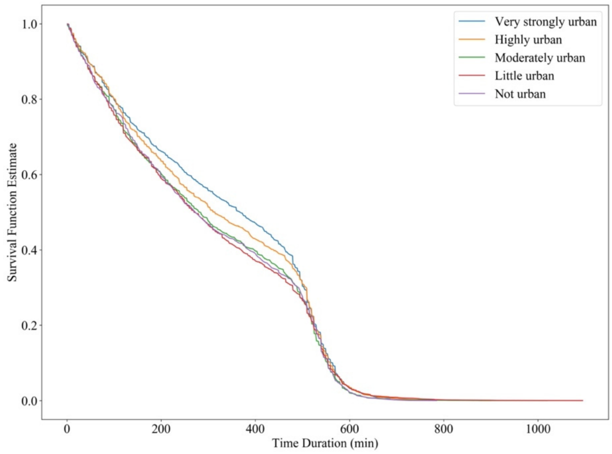


(**b**)



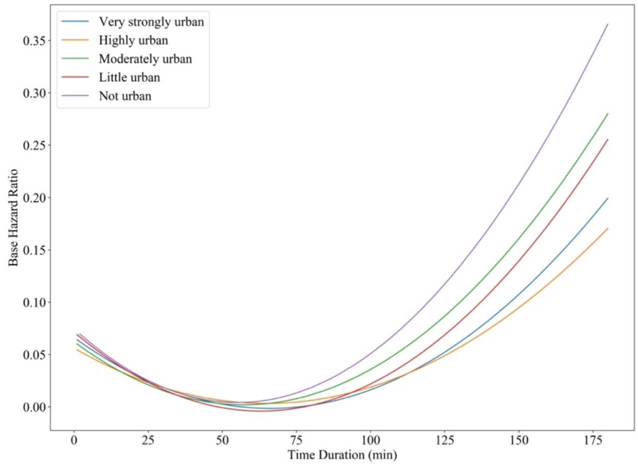
(**c**)

Figure 3. *Cont*.

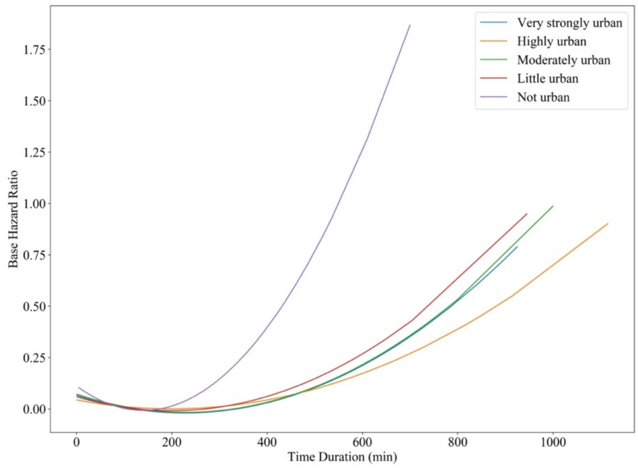


(**d**)

Figure 3. Survival function estimation for different activities and urbanized areas. (**a**) Shopping; (**b**) Social; (**c**) Hiking; (**d**) Work.

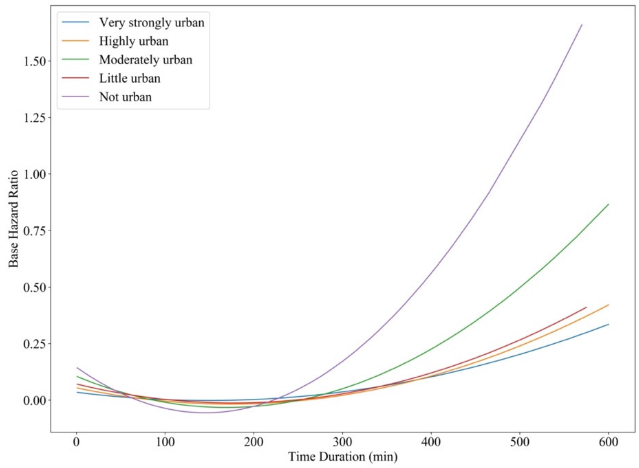


(**a**)

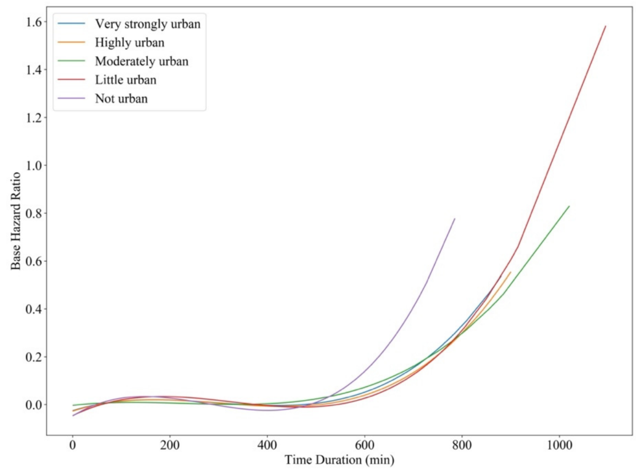


(**b**)

Figure 4. *Cont*.



(**c**)



(**d**)

Figure 4. Fitted hazard curves for different activities and urbanized areas. (**a**) Shopping; (**b**) Social; (**c**) Hiking; (**d**) Work.

5.2. Results of Hazard-Based Duration Model

To examine the effects of the built environment and individual/household characteristics (introduced in Table 2) on the activity durations under the COVID-19 pandemic, a hazard-based duration model (e.g., the Cox proportional hazard model) was calibrated in this study. Note that, different from previous studies that calibrated the models for different groups, we include an urban class variable to consider the effect of urbanization on activity durations. In addition, based on the z-statistics, variables that were significant at a 10% level in one of the four activities were included in the final model. The detailed results are discussed in subsequent sections.

5.2.1. Individual/Household Characteristics

Table 3 represents the estimation results of the socio-demographic variable part. First of all, age tends to significantly influence the duration of social-related activities under the COVID-19 pandemic. The results reveal that, relative to the 18–24 age group, older persons tend to have shorter durations of social-related activities, and the magnitudes of this factor are distributed relatively equally across different age groups. This is probably because people try to avoid non-necessary social activities under the pandemic in order to protect themselves from infection. In particular, the fact that the magnitude for the older group (0.19) is higher than the magnitude for the middle-aged group (0.16) suggests that older people are more concerned by the pandemic, which is fairly consistent with most previous studies [2]. The results also show that young people are more likely to work longer in the pandemic, while elderly people work less. These findings are understandable because the 25–45 age group generally has a strong sense of professionalism, e.g., high level of commitment and motivation to work. In contrast, the elderly tend to work from home and shorten their working hours.

Regarding gender, results show that females are more likely to have a longer shopping activity duration but a shorter social and work-related activity duration. This may indicate that under the COVID-19 pandemic, females have a relatively strong willingness to shop. Unlike males' shopping activities aiming simply to buy things, females prefer to treat it as a hobby and enjoy it [76]. Results also show that females are more reluctant to take part in social and work-related activities under the pandemic. Previous studies also indicated that the reduction of work and social activities decreases the probability of exposure to infected individuals [33]. One possible explanation for this may be that females need to devote more time to their families and take care of their children because of the closure (or the restricted use) of daycare centers and schools.

The results also reveal that people who have a valid driver's license are more likely to spend more time on shopping and work activities. This finding highlights a shift that people expect to need to avoid crowds more than usual to reduce their exposure to COVID-19 [77]. One reason may be that it is convenient for people with a valid driver's license to take part in these activities since they can use a relatively safe and private mode of transport (e.g., car) in their journey. However, the number of cars in the household has an opposite influence on hiking activity duration. It is shown that car-owning households tend to reduce their duration of hiking activities under the COVID-19 pandemic. This may be because people abandoned long-distance hiking in favor of walking or biking around their home locations due to the restrictions on non-essential travel.

Moreover, education level is also found to closely correlate with the activity duration under the COVID-19 pandemic. In particular, we found that people with lower education (e.g., primary and low vocational education) tend to have a longer shopping activity duration. This corresponds well with previous studies [78]. Particularly, relevant studies have also demonstrated that older and less educated people took more time shopping in brick-and-mortar stores [78]. In contrast, all education levels negatively affect the duration of social-related activities, and the effects are relatively stable. This finding confirms that a relatively large proportion of older adults with lower education are less likely to participate in social activities due to physical conditions [79]. In addition, this result may be caused by the policy of home quarantine. In recent years, the policy of home quarantine led to significant changes in the Dutch shopping pattern, and online shopping has become an appealing option to avoid crowded contact in offline malls.

The migration background leads to some interesting interpretations. The effects on shopping and work activity duration are significantly negative, which implies that individuals with migration backgrounds tend to spend more time on shopping and work compared with Dutch people. In addition, people with oriental immigrant backgrounds tend to reduce their hiking activity duration. This is perhaps because the perceived risk of infection for Eastern people is higher than that of people with Western backgrounds [80].

Furthermore, it is found that the coefficients of the income level are significant for social-related activities, indicating that high-income people are more likely to reduce the duration of social activities. This result is in line with the COVID-19 countermeasures urging the public to avoid social interactions outside of the household. In addition, due to the pandemic, people are more inclined to contact each other online unless face-to-face communication is necessary.

Table 3. Estimates for the socio-demographic variables.

	Shopping		Social		Hiking		Work	
	Coef.	**S.E.**	**Coef.**	**S.E.**	**Coef.**	**S.E.**	**Coef.**	**S.E.**
People aged 25 to 45	-	-	0.16 [c]	0.05	-	-	−0.12 [c]	0.04
People aged 46 to 65	-	-	0.17 [c]	0.05	-	-	-	-
People aged 66 and older	-	-	0.19 [c]	0.05	-	-	0.54 [c]	0.06
Female	−0.12 [c]	0.02	0.11 [c]	0.03	-	-	0.09 [c]	0.02
Driver’s license	−0.08 [c]	0.03	-	-	-	-	−0.15 [c]	0.04
Primary education	−0.24 [b]	0.12	0.48 [c]	0.16	-	-	-	-
Lower vocational education	−0.18 [a]	0.11	0.38 [c]	0.15	-	-	-	-
Secondary vocational education	-	-	0.38 [c]	0.15	-	-	-	-
Higher vocational education	-	-	0.4 [c]	0.15	-	-	-	-
Western migration background	−0.07 [b]	0.03	-	-	-	-	−0.08 [b]	0.04
Non-Western migration background	−0.14 [c]	0.04	-	-	0.13 [a]	0.08	−0.08 [b]	0.04
From 100% to 105% low-income threshold	-	-	0.53 [b]	0.22	-	-	-	-
From 106% to 110% low-income threshold	-	-	0.33 [a]	0.2	-	-	-	-
Couple	-	-	-	-	-	-	0.07 [b]	0.03
Couple + child(ren)	-	-	-	-	-	-	0.08 [b]	0.04
Couple + child(ren) + other(s)	-	-	-	-	-	-	0.27 [b]	0.12
Single-parent household + child(ren)	-	-	0.12 [a]	0.06	−0.17 [b]	0.09		
Other household type	-	-	0.42 [b]	0.18	-	-		
Number of people younger than 6 years old	-	-	-	-	−0.27 [a]	0.16	-	-
Number of cars in the household	-	-	-	-	0.36 [a]	0.22	-	-

Note: Only significant values presented; [a]: Significant at the 0.1 level; [b]: Significant at the 0.05 level; [c]: Significant at the 0.01 level.

The results also show that household composition is closely related to activity duration, as one can expect that families with children tend to have shorter social and work-related activity duration. In other words, parents need to shorten their work time and spend more time at home with their children under the pandemic. Moreover, a single-parent household with children is positively and significantly associated with the hiking activity duration, which perhaps means single-parent households have a stronger need for hiking activities during the pandemic.

Similarly, the number of family members younger than six years old has a positive effect on hiking activity duration. As discussed earlier, children have to stay at home due to the closure of childcare facilities but outdoor activities are still necessary. Based on the policy adopted in the Netherlands, there are no restrictions on the travel of underage children. Consequently, parents would prefer to take their children to open green spaces.

5.2.2. Built Environment Characteristics

Regarding the green space variables, results reveal that individuals are more motivated to work and increase their working hours if large-scale green spaces (especially allotment gardens) exist around the workplace. This finding is consistent with our shared knowledge and experience because green space around offices could contribute to physical and mental health. In particular, green space areas around offices could make the office staff exercise properly in their spare time and make them feel energetic and relaxed. However, the result for time spent on socializing in sport and recreational areas is quite the opposite. This result confirms that people are more inclined to socialize in leisure and entertainment venues [81]. Private and quiet recreational areas are more suitable for social activities than noisy sports areas. It is interesting to note that business park and day recreation areas have a negative impact on the shopping activity duration. This reflects that, during the COVID-19 pandemic, shopping desires could be easily satisfied if many other recreational options were available in residential areas.

Similarly, the neighborhood address density (EAD) was employed as a residence measure for each neighborhood. A larger EAD represents a higher density of human settlement. The results show that individuals living in neighborhoods with high EAS have shorter durations for shopping activity. This finding seems logical, as the distribution of shopping malls in large population centers is generally broad, bringing convenience to local people and therefore shortening the shopping duration. From another point of view, EAD, to a certain extent, could reflect the economic status of a family. Individuals who live in villas with a low EAD or communities with large housing areas tend to have a high income [82]. Due to the different economic conditions, as expected, the frequency and duration of shopping activities also differ.

Particular attention should be given to the effects of facility variables. This part will be analyzed in the order of activities: (1) regarding the shopping-related activities, the results show that the duration of shopping activity will increase as the average distance from the center of the neighborhood to the nearest supermarket and daily goods stores increases. This can be attributed to the fact that individuals tend to buy more daily necessities at once to reduce the number of unnecessary trips during the pandemic. (2) For the social and hiking-related activities, it can be found that the duration of social-related activities increases by the increase in the average distance to supermarkets and cinemas. This may be because a long distance to supermarkets and cinemas indirectly indicates the area is less urbanized, and people need to travel for a long time to enjoy social activities. Consequently, when people conduct a social activity, they still tend to take part in it for a long duration even travel cost and risk of COVID-19 infection are high. As shown in Table 4, similar results are also found in the hiking-related activities. When respondents choose to take part in hiking activity, duration increases as the distance to cafés and cafeterias increases. This finding is indeed consistent with our shared knowledge and experience because people need to spend more time finding support facilities during their hiking activities. (3) Lastly, the results reveal that when the number of daycare facilities and restaurants in the neighborhood increases, the duration of work activity also increases. Quite the opposite, the average distance to an important transfer station and main highway entrance is significantly and negatively associated with the duration of work-related activities. The possible reason for these results may be that placing children in nurseries could effectively save parents' time. On the other hand, due to long commuting distance or time, people's intention to work in person will be severely weakened. Furthermore, as the distance from train stations increases, the noise in the community can be effectively reduced, which is beneficial for the workers' productivity and work time [83].

The lower-level land-use variables (e.g., retail and hospitality area, agricultural land, recreational area, residential area, and socio-cultural area) also significantly affect shopping and work-related activities, which is consistent with most previous studies. The negative coefficient of retail and hospitality areas for shopping activity indicates a reduced risk of these activities coming to an end. This suggests that a greater availability of shopping options has a positive influence on the duration of shopping activities. Quite the opposite, the agricultural area significantly and negatively influences shopping time. This result is understandable because a large agricultural area usually means a less developed urban area, and residents have fewer shopping options. In addition, the results show that the larger the recreational and residential areas in a neighborhood, the shorter the in-office work time. One logical explanation for this may be that the probability of exposure to the virus also increases due to the high population density in entertainment venues and residential areas. The negative impact of the soundness and expansion of entertainment venues on work should also be considered.

Table 4. Estimates for the built environment variables.

	Shopping		Social		Hiking		Work	
	Coef.	**S.E.**	**Coef.**	**S.E.**	**Coef.**	**S.E.**	**Coef.**	**S.E.**
Green space variables								
Business park	0.28 [b]	0.11	-	-	-	-	−0.43 [c]	0.15
Day recreation area	0.26 [a]	0.15	-	-	-	-	−2.57 [c]	1.05
Sports area	-	-	0.36 [b]	0.15	-	-	−2.42 [b]	1.16
Allotment garden	-	-	-	-	−0.39 [b]	0.18	−6.11 [b]	2.86
Park	-	-	-	-	-	-	−4.13 [b]	1.77
Recreational area	-	-	−0.57 [b]	0.26	-	-	−4.81 [b]	2.02
Density								
EAD	0.36 [a]	0.21	-	-	-	-	-	-
Facility variables								
DTLS	−0.46 [b]	0.23	−0.55 [a]	0.31	-	-	-	-
DTTS	−0.25 [b]	0.13	-	-	-	-	−0.42 [c]	0.15
DTITS	-	-	-	-	-	-	0.35 [c]	0.14
DTMHE	-	-	-	-	-	-	1.14 [c]	0.39
DTCINEMA	-	-	−0.49 [a]	0.27	-	-	-	-
DTCAFE	-	-	-	-	−0.28 [a]	0.16	-	-
DTCAFETERIA	-	-	-	-	−0.37 [a]	0.22	-	-
NOCINEMAS	0.42 [c]	0.13	-	-	-	-	-	-
NOSTORES	−0.32 [b]	0.14	-	-	-	-	-	-
NOREST	-	-	-	-	-	-	−1.7 [c]	0.58
NOOFSC	-	-	-	-	-	-	−0.54 [a]	0.29
Land use								
Retail and hospitality area	−0.2 [a]	0.11	-	-	-	-	-	-
Total agricultural land	0.17 [a]	0.09	-	-	-	-	-	-
Total recreational area	-	-	-	-	-	-	7.23 [b]	3.25
Residential area	-	-	-	-	-	-	0.41 [b]	0.2
Socio-cultural area	-	-	-	-	-	-	−0.31 [b]	0.16
Urban density								
Urban class_2	−0.08 [c]	0.03	-	-	-	-	-	-
Urban class_3	−0.08 [b]	0.04	-	-	-	-	-	-
Urban class_4	-	-	-	-	−0.21 [c]	0.08	-	-
COVID-19 countermeasures								
Entertainment open	−0.05 [b]	0.03	-	-	-	-	-	-
Face mask requirement	-	-	0.09 [a]	0.05	0.11 [b]	0.05	0.07 [b]	0.04
Two-way interaction								
Entertainment open * NOCINEMAS	−0.25 [b]	0.1	−0.99 [b]	0.47	-	-	-	-
Entertainment open * NOREST	-	-	-	-	-	-	1.58 [c]	0.53

Note: Only significant values presented; [a]: Significant at the 0.1 level; [b]: Significant at the 0.05 level; [c]: Significant at the 0.01 level; Urban class_2 = Highly urban; Urban class_3 = Moderately urban; Urban class_4 = Slightly urban; * means the interaction effect between two variables.

To further investigate the effects of urban class on the duration of different activities, a more detailed estimation (e.g., a quantitative analysis) was conducted to further analyze the correlations. Results are shown in Table 4. It is found that the effects on shopping activity duration in highly and moderately urbanized areas are both positive, indicating that residents in these areas exhibit longer durations of shopping activities compared to individuals in the very strongly urban area. This finding aligns with expectations, as the availability of shopping malls in highly and moderately urbanized areas is comparatively limited compared to very strongly urban areas. Residents in highly and moderately urbanized areas may face less convenience compared to those in very urban areas, leading them to make larger purchases in one go to meet their future needs. In addition, the relatively lower population density in highly and moderately urbanized areas, in contrast to very urban areas, contributes to a reduced risk of COVID-19 infection. As a result, local residents in these areas are more inclined to engage in longer shopping durations. This observation is further supported by the estimated parameter of the variable EAD.

It is interesting to note that a significantly positive correlation is found between the hiking activity duration and the fourth class of urban density (i.e., slightly urbanized area). This may be attributed to the fact that people who live in slightly urbanized areas tend to have a relatively slow-paced lifestyle and fewer options for leisure activities, especially under the COVID-19 pandemic with strict control measures [84]. In addition, due to the availability of large open spaces, the slightly urbanized areas are also more attractive for travelers to take part in hiking activities.

5.2.3. COVID-19 Countermeasures and Interaction Effects

For the effects of COVID-19 countermeasures, it can be found that the opening rule of entertainment venues is positively associated with the duration of shopping activity. This means that the issue of the opening rule of entertainment venues, initially aiming to reduce the shopping demand, tends to increase the shopping duration instead. A plausible explanation for this finding is the shorter opening hours in the pandemic and the limited store attendants. Compared with the pre-pandemic period, customers need to wait a longer time for services under the pandemic (e.g., waiting at the checkouts of supermarkets). In contrast to the positive impact of the opening rule of entertainment venues, the face mask requirement has a negative effect on the duration of social, hiking, and work-related activities. This finding is understandable because the public in Western countries has a perception that wearing a face mask in public open spaces is sometimes considered unusual for a healthy person. In addition, under the pandemic, the enforcement of the face mask requirement signifies a high risk of COVID-19 infection for travelers. This heightened risk not only leads individuals to instinctively shorten their activity duration but also reduce the frequency of their trips to protect themselves from infection. This finding is also supported by relevant literature [11,12].

We also estimated several two-way interactions to identify the interaction effects between built environment factors and COVID-19 countermeasures. Table 4 shows the significant interaction effects between the opening rule of entertainment venues and the number of restaurants and cinemas. It can be found that, when the entertainment venues are allowed to open, the durations of shopping and social-related activities tend to increase with the increasing number of restaurants and cinemas. Quite the opposite, the interaction between the opening rule of entertainment venues and the number of restaurants is significantly and negatively associated with the duration of work-related activities. The possible reason for this may be that people who have been constrained for a long time because of the "intelligent lockdown" are able to re-experience entertainment with the relaxation of measures. Consequently, these palliative measures further contribute towards longer durations of shopping and social-related activities. In addition, due to the long hours of working from home and telecommuting, working time will instead be taken up by other activities as the entertainment venues open up again.

6. Conclusions

Contributing to the literature on investigating the duration of activities, this study examines the effects of various factors on the duration of several typical activities, both in terms of individual characteristics and built environment factors. Unlike earlier studies on normal situations, this study focuses on the effects under the COVID-19 pandemic. More specifically, this paper used a hazard-based duration model to examine the differences in activity duration between different urbanized areas. In particular, the impacts of various COVID-19 countermeasures adopted by the Dutch government, individual characteristics, and local built environment factors on the duration of several representative activities were examined. In this regard, we hope to offer a stepping stone for future study efforts to investigate travel behavior and activity duration concerning the policy and built environments impacts of the COVID-19 or other pandemics in different countries.

The estimation results generally confirm that a relatively large number of factors influence the duration of social and work-related activities. To be more specific, most of these effects are negative, indicating that, under the COVID-19 pandemic, Dutch society as a whole tends to reduce their duration of social and work-related activities. It is also interesting to note that several socio-demographic characteristics and built environment types significantly and positively impact the shopping activity duration. In addition, household composition and green spaces around residential areas are found to correlate closely with the activity duration. As expected, the results also show that driving private vehicles and official COVID-19 countermeasures, to a certain extent, also affect individuals' decision making regarding activity duration. These findings raised a question regarding the coincident objective of introducing COVID-19 countermeasures in many countries. The examination of relevant literature reveals that the primary goal of COVID-19 countermeasures (e.g., non-pharmaceutical interventions) is to minimize mobility and shorten all activities, regardless of their purposes, with the ultimate aim of reducing the transmission of COVID-19 [85]. However, based on the findings of this study, the implemented COVID-19 countermeasures in The Netherlands have only effectively decreased the duration of social and work-related activities. In other words, the COVID-19 countermeasures in the Netherlands have not fully achieved their initial objectives. Despite this, the findings of this study can also serve as a valuable resource for policymakers in developing targeted policies aimed at mitigating diverse types of activities during future health emergencies.

Moreover, the survival curve analysis revealed significant variations in the duration of different activities in terms of both patterns and locations. Specifically, the duration of shopping activities exhibited comparable levels across various urbanized areas. However, a more focused analysis indicated that individuals residing in highly urban and moderately urban areas tend to have longer shopping durations. Consequently, the cumulative impact of these subtle differences significantly influences the overall outcome. Additionally, individuals in highly urbanized areas demonstrated relatively longer durations of work activities compared to those in other urbanized areas. On the contrary, in less urbanized areas, the survival curve illustrating the duration of hiking activities exhibited a relatively rapid decline. This trend can be attributed to factors such as the presence of crop areas and inadequate infrastructural facilities. Subsequent analyses revealed no significant variations in the duration of work activities across different urbanized areas, and only slightly urban areas exhibited a positive impact on hiking activity duration. These findings emphasize the significance of policymakers' acknowledgement of the variations observed in different activities, as well as the often neglected spatial heterogeneities in policy effectiveness, as highlighted in the existing literature [5,6,27]. By recognizing these variations, policymakers can develop tailored intervention policies that are specific to different regions, thereby effectively mitigating mobility during the pandemic. Understanding the heterogeneous nature of activity duration can facilitate the identification of activities that necessitate additional interventions, particularly when the efficacy of current policies on those activities diminishes.

Furthermore, it is important to recognize that activity duration during the pandemic exhibits both social and spatial heterogeneity. The findings related to social heterogeneity offer insights for identifying groups that may exhibit resistance to complying with preventive measures. As revealed in this study, older individuals, those with higher levels of education, and married individuals are more susceptible to the impacts of the COVID-19 pandemic. They perceive adhering to measures aimed at shortening their activity duration as an effective strategy to mitigate the risk of COVID-19 infection. In contrast, young and single males demonstrate a higher level of resilience in maintaining their activity duration despite the increased risk of infection. This particular group, characterized by their continued engagement in social activities, may underestimate the risk of COVID-19 and hold strong beliefs regarding personal immunity [5]. Therefore, special attention should be given to this group, and policymakers should consider implementing appropriate measures to minimize their risk of infection.

The findings of this study could also provide implications for resilient neighborhood planning. Office spaces located in close proximity to plentiful green spaces appear to exhibit a higher level of resilience in terms of work-related activities during the COVID-19 pandemic. In contrast, office spaces situated in residential and commercial areas demonstrate a trend of shortened work durations in the absence of work from home policies. This finding implies that, in order to enhance the working efficiency, the establishment of garden-style office environments (i.e., decentralized, away from shopping and leisure areas, and surrounded by abundant green spaces) should be encouraged in the Netherlands. In addition, prior to the pandemic, the accessibility of shopping and social facilities played a significant role in encouraging shopping and leisure activities. However, during the pandemic, neighborhoods with convenient access to retail facilities were more vulnerable to the impacts of pandemic-related policies, leading to an increase in shopping duration due to social distancing restrictions and reduced staff availability. This observation suggests that regions characterized by concentrated commercial areas are more susceptible to the repercussions of the pandemic. Hence, in future urban planning endeavors, adopting a decentralized approach to distribute commercial facilities within neighborhoods could effectively mitigate the potential challenge of prolonged shopping durations resulting from policy constraints.

In short, this paper is the first of a batch of studies investigating activity durations in the context of the COVID-19 pandemic. More specifically, with the objective to provide valuable insight for future related research, this study examined the relationship between socio-demographic characteristics, built environment factors, and COVID-19 countermeasures regarding the duration of different activities and attempted to find a difference in influence among different urbanized areas. The findings of this study could be of value for transportation professionals and health protection agencies to develop transportation plans and establish COVID-19 policies, as well as for future pandemics or any other public health emergencies. Nevertheless, although this research can be regarded as a supplement to the literature on the impacts of the COVID-19 pandemic and the influential factors of activity duration, this topic could be further explored in future studies, especially for the consideration of heterogeneous effects of COVID-19 countermeasures. More narrowly, this paper has not considered the differentiated impacts of COVID-19 countermeasures on different populations. In addition, in the current study, we only examine to what extent the included factors influence the duration of a single activity in a single trip. However, a trip may have several different activities. Thus, it would be interesting in future research to examine the combined influences on sequential activities. Another limitation of this research is the data. Indeed, the data used in this study were obtained from the self-reported national travel survey of The Netherlands in 2020 and have not been objectively confirmed, e.g., by GPS tracking, which may have led to biased estimation. Therefore, the results could be more convincing if more accurate data (e.g., GPS data) could be collected in future studies. Lastly, it is important to note that the scope of this study is limited to investigating the determinants of activity duration during the pandemic. Future research endeavors could be intriguingly expanded to explore the underlying determinants that influence the differences in activity durations between the pre-pandemic and pandemic periods, provided the availability of suitable data.

Author Contributions: Conceptualization, C.L. and M.S.; methodology, X.G.; software, X.Z.; validation, M.S.; investigation, C.L.; resources, M.S.; data curation, X.Z. and X.G.; writing—original draft, C.L.; writing—review & editing, C.C.; supervision, C.C.; project administration, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by China Scholarship Council grant number 201906060030.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data employed in this study are available from the Central Bureau of Statistics (CBS), The Netherlands. The data are open access, and the authors are authorized to freely use, copy, and redistribute the material in any medium or format. Therefore, the data are available directly from the website of CBS Open Data StatLine or the authors if requested.

Acknowledgments: This paper was jointly supported by the Science Research Development Program of Higher Education of the China Association of Higher Education (No. 22DL0403) and the Key R&D Project (Soft Science) of Shandong Province, China (No. 2021RZA02025). The authors would also like to thank the China Scholarship Council (CSC, No. 201906060030) for support to Chao Chen.

Conflicts of Interest: The authors declare that they have no competing interest that might have had an effect on the work reported in the study.

References

1. Statistics on the Hopkins Outbreak in the USA. TRACKING: Follow Global Cases and Trends, Updated Daily. 2021. Available online: https://coronavirus.jhu.edu/data (accessed on 3 October 2021).
2. Shelat, S.; Cats, O.; van Cranenburgh, S. Avoiding the Crowd: Traveller Behaviour in Public Transport in the Age of COVID-19. *arXiv* **2021**, arXiv:2104.10973.
3. Abdullah, M.; Dias, C.; Muley, D.; Shahin, M. Exploring the impacts of COVID-19 on travel behavior and mode preferences. *Transp. Res. Interdiscip. Perspect.* **2020**, *8*, 100255. [CrossRef] [PubMed]
4. Fatmi, M.R.; Thirkell, C.; Hossain, S. COVID-19 and Travel: How Our Out-of-home Travel Activity, In-home Activity, and Long-Distance Travel Have Changed. *Transp. Res. Interdiscip. Perspect.* **2021**, *10*, 100350. [CrossRef] [PubMed]
5. Liu, C.; Zhang, W. Social and spatial heterogeneities in COVID-19 impacts on individual's metro use: A big-data driven causality inference. *Appl. Geogr.* **2023**, *155*, 102947. [CrossRef] [PubMed]
6. Zhang, W.; Li, J. A quasi-experimental analysis on the causal effects of COVID-19 on urban park visits: The role of park features and the surrounding built environment. *Urban For. Urban Green.* **2023**, *82*, 127898. [CrossRef] [PubMed]
7. Elvén, M.; Kerstis, B.; Stier, J.; Hellström, C.; von Heideken Wågert, P.; Dahlen, M.; Lindberg, D. Changes in physical activity and sedentary behavior before and during the COVID-19 pandemic: A Swedish population study. *Int. J. Environ. Res. Public Health* **2022**, *19*, 2558. [CrossRef]
8. Rafiq, R.; McNally, M.G.; Uddin, Y.S.; Ahmed, T. Impact of working from home on activity-travel behavior during the COVID-19 Pandemic: An aggregate structural analysis. *Transp. Res. Part A Policy Pract.* **2022**, *159*, 35–54. [CrossRef]
9. Stokes, E.C.; Román, M.O. Tracking COVID-19 urban activity changes in the Middle East from nighttime lights. *Sci. Rep.* **2022**, *12*, 8096. [CrossRef]
10. Batur, I.; Dirks, A.C.; Bhat, C.R.; Polzin, S.E.; Chen, C.; Pendyala, R.M. Analysis of changes in time use and activity participation in response to the COVID-19 pandemic in the United States: Implications for well-being. *Transp. Res. Rec. J. Transp. Res. Board* **2023**. [CrossRef]
11. Chen, C.; Feng, T.; Gu, X. Role of latent factors and public policies in travel decisions under COVID-19 pandemic: Findings of a hybrid choice model. *Sustain. Cities Soc.* **2022**, *78*, 103601. [CrossRef]
12. Chen, C.; Feng, T.; Gu, X.; Yao, B. Investigating the effectiveness of COVID-19 pandemic countermeasures on the use of public transport: A case study of The Netherlands. *Transp. Policy* **2022**, *117*, 98–107. [CrossRef]
13. Bedford, J.; Enria, D.; Giesecke, J.; Heymann, D.L.; Ihekweazu, C.; Kobinger, G.; Lane, H.C.; Memish, Z.; Oh, M.-D.; Sall, A.A.; et al. COVID-19: Towards controlling of a pandemic. *Lancet* **2020**, *395*, 1015–1018. [CrossRef]
14. Brooks, S.K.; Webster, R.K.; Smith, L.E.; Woodland, L.; Wessely, S.; Greenberg, N.; Rubin, G.J. The psychological impact of quarantine and how to reduce it: Rapid review of the evidence. *Lancet* **2020**, *395*, 912–920. [CrossRef]
15. Meira, C.M., Jr.; Meneguelli, K.S.; Leopoldo, M.P.; Florindo, A.A. Anxiety and Leisure-Domain Physical Activity Frequency, Duration, and Intensity during COVID-19 Pandemic. *Front. Psychol.* **2020**, *11*, 3758. [CrossRef]
16. Zhao, L.; Shen, Z.; Zhang, Y.; Sheng, F. Study on the Impact of the Objective Characteristics and Subjective Perception of the Built Environment on Residents' Physical Activities in Fuzhou, China. *Sustainability* **2020**, *12*, 329. [CrossRef]
17. Hahm, Y.; Yoon, H.; Choi, Y. The effect of built environments on the walking and shopping behaviors of pedestrians; A study with GPS experiment in Sinchon retail district in Seoul, South Korea. *Cities* **2019**, *89*, 1–13. [CrossRef]
18. Koohsari, M.J.; Shibata, A.; Ishii, K.; Kurosawa, S.; Yasunaga, A.; Hanibuchi, T.; Nakaya, T.; Mavoa, S.; McCormack, G.R.; Oka, K. Built environment correlates of objectively-measured sedentary behaviours in densely-populated areas. *Health Place* **2020**, *66*, 102447. [CrossRef]
19. Pratama, V.; Santoso, I.; A Mustaniroh, S. Development strategy of SMEs in the new normal era of coronavirus disease 2019 (COVID-19): A literature review. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *733*, 012058. [CrossRef]
20. Sá, C.; Dismuke, C.E.; Guimarães, P. Survival analysis and competing risk models of hospital length of stay and discharge destination: The effect of distributional assumptions. *Health Serv. Outcomes Res. Methodol.* **2007**, *7*, 109–124. [CrossRef]
21. Pang, J.; Krathaus, A.; Benedyk, I.; Ahmed, S.S.; Anastasopoulos, P.C. A temporal instability analysis of environmental factors affecting accident occurrences during snow events: The random parameters hazard-based duration model with means and variances heterogeneity. *Anal. Methods Accid. Res.* **2022**, *34*, 100215. [CrossRef]

22. Hasan, S.; Mesa-Arango, R.; Ukkusuri, S. A random-parameter hazard-based model to understand household evacuation timing behavior. *Transp. Res. Part C Emerg. Technol.* **2013**, *27*, 108–116. [CrossRef]
23. Anastasopoulos, P.C.; Bin Islam, M.; Perperidou, D.; Karlaftis, M.G. Hazard-based analysis of travel distance in urban environments: Longitudinal data approach. *J. Urban Plan. Dev.* **2012**, *138*, 53–61. [CrossRef]
24. Berg, P.V.D.; Arentze, T.; Timmermans, H. A latent class accelerated hazard model of social activity duration. *Transp. Res. Part A Policy Pract.* **2012**, *46*, 12–21. [CrossRef]
25. Dissanayake, D. Watching the clock on the way to work? Analysing trends in commuting activities, modes and gender differences in commute times, using hazard-based duration modelling methods. *J. Transp. Geogr.* **2017**, *65*, 188–199. [CrossRef]
26. Hosseini, S.A.R.S.; Hatamzadeh, Y. Walking duration in daily travel: An analysis among males and females using a hazard-based model. *Transp. A Transp. Sci.* **2021**, *17*, 1150–1170. [CrossRef]
27. Zhang, W.; Ning, K. Spatiotemporal heterogeneities in the causal effects of mobility intervention policies during the COVID-19 outbreak: A spatially interrupted time-series (SITS) analysis. *Ann. Assoc. Am. Geogr.* **2023**, *113*, 1112–1134. [CrossRef]
28. Ammar, A.; Brach, M.; Trabelsi, K.; Chtourou, H.; Boukhris, O.; Masmoudi, L.; Bouaziz, B.; Bentlage, E.; How, D.; Ahmed, M.; et al. Effects of COVID-19 home confinement on eating behaviour and physical activity: Results of the ECLB-COVID19 international online survey. *Nutrients* **2020**, *12*, 1583. [CrossRef]
29. Thunström, L.; Newbold, S.C.; Finnoff, D.; Ashworth, M.; Shogren, J.F. The benefits and costs of using social distancing to flatten the curve for COVID-19. *J. Benefit-Cost Anal.* **2020**, *11*, 179–195. [CrossRef]
30. Liu, Z.; Stern, R. Quantifying the Traffic Impacts of the COVID-19 Shutdown. *J. Transp. Eng. Part A Syst.* **2021**, *147*, 04021014. [CrossRef]
31. Park, K.; Chamberlain, B.; Song, Z.; Esfahani, H.N.; Sheen, J.; Larsen, T.; Novack, V.L.; Licon, C.; Christensen, K. A double jeopardy: COVID-19 impacts on the travel behavior and community living of people with disabilities. *Transp. Res. Part A Policy Pract.* **2022**, *156*, 24–35. [CrossRef]
32. Cohen, J.; Kupferschmidt, K. Countries test tactics in 'war' against COVID-19. *Science* **2020**, *367*, 1287–1288. [CrossRef]
33. Engle, S.; Stromme, J.; Zhou, A. Staying at home: Mobility effects of COVID-19. *SSRN Electron. J.* **2020**, 16. [CrossRef]
34. Anderson, R.M.; Heesterbeek, H.; Klinkenberg, D.; Hollingsworth, T.D. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *Lancet* **2020**, *395*, 931–934. [CrossRef]
35. de Jonge, E.; Kloppenburg, R.; Hendriks, P. The impact of the COVID-19 pandemic on social work education and practice in the Netherlands. *Soc. Work. Educ.* **2020**, *39*, 1027–1036. [CrossRef]
36. van Leeuwen, M.; Klerks, Y.; Bargeman, B.; Heslinga, J.; Bastiaansen, M. Leisure will not be locked down–insights on leisure and COVID-19 from The Netherlands. *World Leis. J.* **2020**, *62*, 339–343. [CrossRef]
37. Wennekers, A.; Roeters, A.; van den Broek, A.; Pulles, I. Vrije Tijd. De Sociale Staat van Nederland: 2018. 2018. Available online: https://digitaal.scp.nl/ssn2018/vrije-tijd (accessed on 4 May 2023).
38. Li, L.; Li, X.; Niu, N.; He, J. Uneven impacts of COVID-19 on residents' utilization of urban parks: A case study of Guangzhou, China. *Appl. Geogr.* **2023**, *153*, 102905. [CrossRef]
39. Susilo, Y.O.; Axhausen, K.W. Repetitions in individual daily activity–travel–location patterns: A study using the Herfindahl–Hirschman Index. *Transportation* **2014**, *41*, 995–1011. [CrossRef]
40. Pinjari, A.; Bhat, C. A multiple discrete–continuous nested extreme value (MDCNEV) model: Formulation and application to non-worker activity time-use and timing behavior on weekdays. *Transp. Res. Part B Methodol.* **2010**, *44*, 562–583. [CrossRef]
41. Bláfoss, R.; Micheletti, J.K.; Sundstrup, E.; Jakobsen, M.D.; Bay, H.; Andersen, L.L. Is fatigue after work a barrier for leisure-time physical activity? Cross-sectional study among 10,000 adults from the general working population. *Scand. J. Public Health* **2019**, *47*, 383–391. [CrossRef]
42. Sreela, P.; Melayil, S.; Anjaneyulu, M. Modeling of Shopping Participation and Duration of Workers in Calicut. *Procedia-Soc. Behav. Sci.* **2013**, *104*, 543–552. [CrossRef]
43. Sarangi, P.; Manoj, M. Analysis of activity participation and time use decisions of partners: The context of low-and high-income households. *Transportation* **2021**, *49*, 1017–1058. [CrossRef]
44. Dane, G.; Arentze, T.A.; Timmermans, H.J.; Ettema, D. Simultaneous modeling of individuals' duration and expenditure decisions in out-of-home leisure activities. *Transp. Res. Part A Policy Pract.* **2014**, *70*, 93–103. [CrossRef]
45. Clark, B.; Chatterjee, K.; Melia, S. Changes to commute mode: The role of life events, spatial context and environmental attitude. *Transp. Res. Part A Policy Pract.* **2016**, *89*, 89–105. [CrossRef]
46. Wang, Z.; Ettema, D.; Helbich, M. Objective environmental exposures correlate differently with recreational and transportation walking: A cross-sectional national study in The Netherlands. *Environ. Res.* **2021**, *194*, 110591. [CrossRef]
47. Colom, A.; Mavoa, S.; Ruiz, M.; Wärnberg, J.; Muncunill, J.; Konieczna, J.; Vich, G.; Barón-López, F.J.; Fitó, M.; Salas-Salvadó, J.; et al. Neighbourhood walkability and physical activity: Moderating role of a physical activity intervention in overweight and obese older adults with metabolic syndrome. *Age Ageing* **2021**, *50*, 963–968. [CrossRef]
48. Vich, G.; Delclòs-Alió, X.; Maciejewska, M.; Marquet, O.; Schipperijn, J.; Miralles-Guasch, C. Contribution of park visits to daily physical activity levels among older adults: Evidence using GPS and accelerometery data. *Urban For. Urban Green.* **2021**, *63*, 127225. [CrossRef]
49. Palma, D.; Enam, A.; Hess, S.; Calastri, C.; Sourd, R.C.D. Modelling multiple occurrences of activities during a day: An extension of the MDCEV model. *Transp. B Transp. Dyn.* **2021**, *9*, 456–478. [CrossRef]

50. Bhat, C.R. The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transp. Res. Part B Methodol.* **2008**, *42*, 274–303. [CrossRef]
51. Enam, A.; Konduri, K.C.; Pinjari, A.R.; Eluru, N. An integrated choice and latent variable model for multiple discrete continuous choice kernels: Application exploring the association between day level moods and discretionary activity engagement choices. *J. Choice Model.* **2018**, *26*, 80–100. [CrossRef]
52. Calastri, C.; Hess, S.; Daly, A.; Carrasco, J.A. Does the social context help with understanding and predicting the choice of activity type and duration? An application of the Multiple Discrete-Continuous Nested Extreme Value model to activity diary data. *Transp. Res. Part A Policy Pract.* **2017**, *104*, 1–20. [CrossRef]
53. Saxena, S.; Pinjari, A.R.; Bhat, C.R.; Mondal, A. A flexible multiple discrete–continuous probit (MDCP) model: Application to analysis of expenditure patterns of domestic tourists in India. *Transportation* **2023**, 1–28. [CrossRef]
54. Baloi, D.; Price, A.D. Modelling global risk factors affecting construction cost performance. *Int. J. Proj. Manag.* **2003**, *21*, 261–269. [CrossRef]
55. Salah, A.; Moselhi, O. Contingency modelling for construction projects using fuzzy-set theory. *Eng. Constr. Arch. Manag.* **2015**, *22*, 214–241. [CrossRef]
56. Zadeh, L.A. Fuzzy sets. *Inf. Control.* **1965**, *8*, 338–353. [CrossRef]
57. Gładysz, B. Fuzzy-probabilistic PERT. *Ann. Oper. Res.* **2017**, *258*, 437–452. [CrossRef]
58. Gao, J.; Kamphuis, C.B.M.; Dijst, M.; Helbich, M. The role of the natural and built environment in cycling duration in The Netherlands. *Int. J. Behav. Nutr. Phys. Act.* **2018**, *15*, 82. [CrossRef]
59. Cox, D.R. Regression models and life-tables. *J. R. Stat. Soc. Ser. B (Methodol.)* **1972**, *34*, 187–202. [CrossRef]
60. Yee, J.L.; A Niemeier, D. Analysis of activity duration using the Puget sound transportation panel. *Transp. Res. Part A Policy Pract.* **2000**, *34*, 607–624. [CrossRef]
61. Anastasopoulos, P.C.; Fountas, G.; Sarwar, T.; Karlaftis, M.G.; Sadek, A.W. Transport habits of travelers using new energy type modes: A random parameters hazard-based approach of travel distance. *Transp. Res. Part C Emerg. Technol.* **2017**, *77*, 516–528. [CrossRef]
62. Centraal Bureau voor de Statistiek (CBS); Rijkswaterstaat (RWS-WVL) (2021a): Onderzoek Onderweg in Nederland—ODiN 2020. DANS. Available online: https://www.cbs.nl/nl-nl/longread/rapportages/2021/onderweg-in-nederland{-}{-}odin{-}{-}-2020-onderzoeksbeschrijving?onepage=true (accessed on 4 May 2023).
63. Centraal Bureau Voor de Statistiek (CBS), Bodemgebruik, Wijk- en Buurtcijfers 2015. 2021. Available online: https://opendata.cbs.nl/statline/portal.html?_la=nl&_catalog=CBS&tableId=84324NED (accessed on 4 May 2023).
64. Centraal Bureau voor de Statistiek (CBS). Nabijheid Voorzieningen; Afstand Locatie, Wijk- en Buurtcijfers 2019. 2021. Available online: https://opendata.cbs.nl/statline/portal.html?_la=nl&_catalog=CBS&tableId=84718NED (accessed on 4 May 2023).
65. Centraal Bureau voor de Statistiek (CBS). Kerncijfers Wijken en Buurten 2020. 2021. Available online: https://opendata.cbs.nl/statline/portal.html?_la=nl&_catalog=CBS&tableId=84799NED (accessed on 4 May 2023).
66. Centraal Bureau voor de Statistiek (CBS). Nabijheid Voorzieningen; Afstand Locatie, Wijk- en Buurtcijfers 2020. 2021. Available online: https://opendata.cbs.nl/statline/portal.html?_la=nl&_catalog=CBS&tableId=84953NED (accessed on 4 May 2023).
67. Zagorskas, J. GIS-based Modelling and estimation of land use mix in urban environment. *Int. J. Educ. Learn. Syst.* **2016**, *1*, 284–293.
68. Eisenmann, C.; Nobis, C.; Kolarova, V.; Lenz, B.; Winkler, C. Transport mode use during the COVID-19 lockdown period in Germany: The car became more important, public transport lost ground. *Transp. Policy* **2021**, *103*, 60–67. [CrossRef] [PubMed]
69. Hensher, D.A.; Mannering, F.L. Hazard-based duration models and their application to transport analysis. *Transp. Rev.* **1994**, *14*, 63–82. [CrossRef]
70. Kaplan, E.L.; Meier, P. Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* **1958**, *53*, 457–481. [CrossRef]
71. Mohammadian, A.; Doherty, S.T. Modeling activity scheduling time horizon: Duration of time between planning and execution of pre-planned activities. *Transp. Res. Part A Policy Pract.* **2006**, *40*, 475–490. [CrossRef]
72. Dhoke, A.; Kumar, A.; Ghosh, I. Hazard-Based Duration Approach to Pedestrian Crossing Behavior at Signalized Intersections. *Transp. Res. Rec. J. Transp. Res. Board* **2021**, *2675*, 519–532. [CrossRef]
73. Bhat, C. Duration modeling: A methodological review with applications in activity-travel analysis. In Proceedings of the 80th Annual Meeting of the Transportation Research Board, Washington, DC, USA, 7–11 January 2001.
74. Liu, B.; Chen, H.; Gan, X. How much is too much? The influence of work hours on social development: An empirical analysis for OECD countries. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4914. [CrossRef]
75. Inggamara, A.; Pierewan, A.C.; Ayriza, Y. Work–life balance and social support: The influence on work engagement in the Sixth European Working Conditions Survey. *J. Employ. Couns.* **2022**, *59*, 17–26.
76. Hatamzadeh, Y.; Habibian, M.; Khodaii, A. Walking mode choice across genders for purposes of work and shopping: A case study of an Iranian city. *Int. J. Sustain. Transp.* **2020**, *14*, 389–402. [CrossRef]
77. Tirachini, A.; Cats, O. COVID-19 and public transportation: Current assessment, prospects, and research needs. *J. Public Transp.* **2020**, *22*, 1. [CrossRef]
78. Farag, S.; Dijst, M.; Lanzendorf, M. Exploring the use of e-shopping and its impact on personal travel behavior in The Netherlands. *Transp. Res. Rec. J. Transp. Res. Board* **2003**, *1858*, 47–54. [CrossRef]

79. Evertsson, M.; England, P.; Mooi-Reci, I.; Hermsen, J.; de Bruijn, J.; Cotter, D. Is gender inequality greater at lower or higher educational levels? Common patterns in The Netherlands, Sweden, and the United States. *Soc. Politi-Int. Stud. Gender State Soc.* **2009**, *16*, 210–241. [CrossRef]
80. Zha, J.; Tan, T.; Ma, S.; He, L.; Filimonau, V. Exploring tourist opinion expression on COVID-19 and policy response to the pandemic's occurrence through a content analysis of an online petition platform. *Curr. Issues Tour.* **2021**, *25*, 261–286. [CrossRef]
81. Tsaligopoulos, A.; Kyvelou, S.; Votsi, N.-E.; Karapostoli, A.; Economou, C.; Matsinos, Y.G. Revisiting the Concept of Quietness in the Urban Environment—Towards Ecosystems' Health and Human Well-Being. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3151. [CrossRef]
82. Rong, P.; Zhang, Y.; Qin, Y.; Liu, G.; Liu, R. Spatial differentiation of carbon emissions from residential energy consumption: A case study in Kaifeng, China. *J. Environ. Manag.* **2020**, *271*, 110895. [CrossRef]
83. Grubliauskas, R.; Strukčinskienė, B.; Raistenskis, J.; Strukčinskaitė, V.; Buckus, R.; Januševičius, T.; da Silva Pereira, P.A. Effects of urban rail noise level in a residential area. *J. Vibroeng.* **2014**, *16*, 987–996.
84. Pullano, G.; Valdano, E.; Scarpa, N.; Rubrichi, S.; Colizza, V. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: A population-based study. *Lancet Digit. Health* **2020**, *2*, e638–e649. [CrossRef]
85. Tully, M.A.; McMaw, L.; Adlakha, D.; Blair, N.; McAneney, J.; McAneney, H.; Carmichael, C.; Cunningham, C.; Armstrong, N.C.; Smith, L. The effect of different COVID-19 public health restrictions on mobility: A systematic review. *PLoS ONE* **2021**, *16*, e0260919. [CrossRef]

sustainability

MDPI

Article

Optimal Parking Path Planning and Parking Space Selection Based on the Entropy Power Method and Bayesian Network: A Case Study in an Indoor Parking Lot

Jingwei Xue, Jiaqing Wang *, Jiyang Yi, Yang Wei *, Kaijian Huang, Daming Ge and Ruiyu Sun

College of Civil Engineering, Nanjing Forestry University, Nanjing 210037, China; xuejingwei@njfu.edu.cn (J.X.); yjyaegean@njfu.edu.cn (J.Y.); huangkaijian@njfu.edu.cn (K.H.); 13812179297@163.com (R.S.)
* Correspondence: jiaqingw@njfu.edu.cn (J.W.); wy78@njfu.edu.cn (Y.W.)

Abstract: According to the vehicle dynamics model and the requirements of reliable safety and minimal time, the path planning problem of parking in different types of parking spaces is solved by obstacle avoidance analysis and motion analysis in the case of the optimal solution, and the parking trajectory from the initial position to the designated parking space is obtained. In the static situation, different parking spaces in the parking space are occupied; analyze the parking space type, parking space left and right occupancy situation, and the distance between the vacant parking space and the starting point location of unoccupied cars; and establish the attribute information matrix R_0 of the vacant parking space and calculate the KMO value of the matrix R_0. This is completed to determine the weak correlation between the attributes of the vacant parking space and use the matrix R_0 as the original evaluation matrix of the entropy weight method, using the entropy weight method to calculate the three attributes of parking space type, parking space left and right occupancy situation, and distance between starting point and parking space. These results are weighted in the optimal parking space selection process, the difficulty score of the vacant parking space is determined, and the optimal parking space is determined through the ranking of the scores. In the dynamic case, the number of parking spaces and parking space usage will change over time, with the help of the Bayesian network, the existing parking spaces and number of spaces in the parking lot at the previous moment are learned according to the computer clock, which can be used to reason about the number of parking spaces and parking space availability in the parking lot at the next moment. The weights of the three attributes of parking space type, parking space left and right situation, and distance between the starting point and parking space are updated in the case of a dynamic change of parking space, and then the parking difficulty score of a new vacant parking space using the entropy weight method is used to select the optimal parking space in the dynamic situation. The optimized parking path planning and parking space selection method could contribute to enhancing parking efficiency for the sustainable management of indoor parking lots.

Keywords: parking management; indoor parking lot; parking path planning; parking area selection; Bayesian network

check for updates

Citation: Xue, J.; Wang, J.; Yi, J.; Wei, Y.; Huang, K.; Ge, D.; Sun, R. Optimal Parking Path Planning and Parking Space Selection Based on the Entropy Power Method and Bayesian Network: A Case Study in an Indoor Parking Lot. *Sustainability* **2023**, *15*, 8450. https://doi.org/10.3390/su15118450

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 20 April 2023
Revised: 16 May 2023
Accepted: 19 May 2023
Published: 23 May 2023

1. Introduction

Over the decades, with the rapid development of economics, personal vehicles have gradually entered a large number of households, becoming the first choice for traveling. With the increasing scale of car ownership in China, the contradiction between the supply and demand of cars and parking spaces in the city is getting bigger and bigger, and the problem of difficulty parking in public parking lots has become particularly prominent.

The difficulty of "parking" comes from two main sources: (1) it is not easy to choose the right parking space, and (2) there is a certain technical difficulty in parking operation. With the limited parking space, the driver's line of sight is obstructed, and it is difficult to quickly and easily find the optimal parking space. The small parking space makes parking a more complex task. In general, drivers park with the assistance of the backup radar, backup camera, etc., but these assists still have limitations. The car may still be in the parking process and two cars scraping or collision with pedestrians can occur. Indoor parking-assisted parking technology has strong practical significance, which can provide practical convenience for human beings; shorten the time required for vehicle parking; significantly improve the safety, stability, and convenience of the parking process; improve the accuracy of vehicle perception of the environment; make the vehicle park accurately into the target parking space; and greatly reduce the risk caused by improper operation, insufficient technology, and emotional and psychological instability of the driver from the operational level; thus, it has broad research prospects and practical significance for project implementation.

Lee [1] et al. proposed an HDL-32E LIDAR automatic parking method based on self-driving cars by improving the random tree (RRT) algorithm for rapid exploration and using a fuzzy logic controller to control the brake and gas pedal for speed stabilization. Zhang [2] et al. built an environment model close to the actual environment, and then constructed a reward function to evaluate and filter the parking data. Finally, a reinforcement learning-based automatic parking method is proposed by using neural networks to learn parking strategies from the filtered data. Gan [3] et al. improved the a-star algorithm by using RS curves and potential functions; introduced an NMPC trajectory optimizer based on vehicle kinematic constraints, minimum deviation objective, and obstacle avoidance objective; and proposed a spatiotemporal heuristic method. Zhang [4] et al. first calibrated the vehicle position, obstacles, and parking spaces, and then generated the kinematic space constraints and proposed a shortest parking path generation strategy for an automatic parking system based on a bidirectional width first-search algorithm and an improved Bellman–Ford algorithm. Su [5] et al. proposed a secondary parallel automatic parking method of endpoint regionalization, and then designed a reasonable parking terminal area, planned a secondary parallel parking path, established a parking path function with constraints, and used a genetic algorithm to optimize the parking path function as the objective function; the design result is 4.1% shorter than the original path. García [6] et al. provided a solution that is based on the analysis of zenith images using artificial vision, and the solution can not only detect whether there are vehicles in the parking space, but also use the region-based convolutional neural network to detect the area occupied by the parking space, with an accuracy rate of 98.21%; finally, the appropriate parking space is specified for the vehicle. Oetiker [7] et al. proposed a closed-loop method to solve the semi-autonomous non-holonomic vehicle parking assistance method. The method allowed for a priori unknown and potential high speed and dynamic environment changes; at the same time, the computational cost of the storage efficient algorithm was very low, and it could well perceive the surrounding environment and adapt, thus gradually improving parking safety. Zhang [8] et al. proposed a DWT-Bi-LSTM parking space availability prediction model based on historical parking data: first, the threshold method was used to process the sequence data; then, the model learned from the historical denoising data to effectively avoid errors; finally, the forward and backward LSTM further improved the prediction accuracy, and the prediction speed was quicker. Peng [9] et al. proposed a new c-non-holonomic trajectory method: first, by defining the c-nonholonomic configuration, the c-non-holonomic trajectory was obtained; second, the Lyapunov method was used to demonstrate the convergence of the globally discontinuous time-invariant feedback controller; finally, the motion trajectory during parking was analyzed.

Although the research [10–15] on assisted parking technology is mostly concentrated in the field of unmanned driving, unmanned vehicles will not be widely popularized in the next decade, so it is more desirable to introduce some small and intelligent algorithms into public life to help them have a faster, better, and safer experience in the daily parking process.

This paper presents a fast-assisted parking method based on the entropy power method and Bayesian network for optimal parking spot selection and path planning for an optimized parking lot operation. The vehicle kinematic model and parking plot were established for analyzing the optimum parking path within a common parking condition, the different directions of the parking path, and the locations of parking spaces. Both the static and dynamic conditions were studied and analyzed based on the proposed models. The results represented an efficient method for parking path guidance and parking space selection, which could be applied in parking lot arrangements for fast and safe parking. With improved parking efficiency, the sustainable management of indoor parking lots could be better facilitated in the future application.

2. Methods and Model Establishment

2.1. Vehicle Model

2.1.1. Vehicle Parameters

The vehicle parameters are as follows: body length of 4.9 m and width of 1.8 m; car wheelbase of 2.8 m and wheel spacing of 1.7 m; maximum acceleration of 3.0 m/s, maximum deceleration limit of -6.0 m/s^2, and acceleration of no more than 20.0 m/s^3; steering wheel maximum angle of 470°, steering wheel and front wheel angle of transmission ratio of 16:1 (steering wheel rotation 16°), maximum steering wheel rotation angle of 470°, ratio of the steering wheel to front wheel rotation of 16:1 (steering wheel rotation 16°, front wheel rotation 1°), and maximum speed of the steering wheel of 400°/s.

2.1.2. Vehicle Kinematic Model

Firstly, the vehicle solid model is simplified into a two-dimensional plane kinematics model, as shown in Figure 1. The points $A(x_A, y_A)$, $B(x_B, y_B)$, $C(x_C, y_C)$, and $D(x_D, y_D)$ in Figure 1 represent the left front vertex, right front vertex, right rear vertex, and left rear vertex of the vehicle, respectively; $b(x_b, y_b)$, $c(x_c, y_c)$, and $d(x_d, y_d)$ represent the contact points between the left front wheel, right front wheel, right rear wheel, and left rear wheel of the vehicle and the ground, respectively; $\left(x_f, y_f\right)$ and (x_r, y_r) represent the center point coordinates of the front axis and the center point coordinates of the rear axis of the vehicle, respectively; v_f and v_r represent the center point velocities of the front axis and the rear axis, respectively. l L indicates the length of the vehicle body, d indicates the width of the vehicle body, δ_i and δ_o respectively indicate the Ackermann angle of the left front wheel and the right front wheel of the vehicle, φ indicates the equivalent Ackermann angle of the two front wheels, and θ indicates the heading angle of the vehicle.

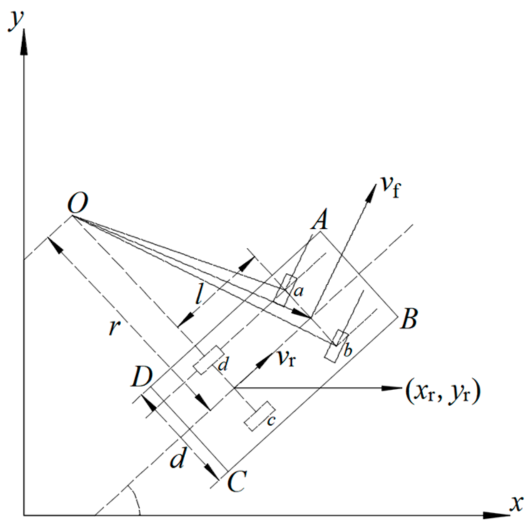


Figure 1. Vehicle kinematic model.

From Figure 1, it can be seen that there is a geometric relationship between the front axle center point $\left(x_f, y_f\right)$ and the rear axle center point (x_r, y_r) of the vehicle, as shown in Equation (2).

$$\begin{cases} x_r = x_f - l \cdot \cos\theta \\ y_r = x_f - l \cdot \cos\theta \end{cases} \quad (1)$$

The differentiation of Equation (2) yields

$$\begin{cases} \dot{x}_r = \dot{x}_f + \dot{\theta} \cdot l \cdot \sin\theta \\ \dot{y}_r = \dot{y}_f - \dot{\theta} \cdot l \cdot \cos\theta \end{cases} \quad (2)$$

In the process of parking the vehicle, considering the safety issue, the vehicle is considered to be in a low-speed motion, at which time the lateral force of the rear wheels of the vehicle is ignored, and cases such as sideslip and sideswipe are not considered. Therefore, the velocity of the rear wheels of the vehicle in the axial direction is 0, and thus the velocity component of the velocity of the rear axle center point in the x-axis and y-axis directions has a certain geometric relationship, as shown in Equation (4).

$$v_{rx} \cdot \sin\theta = v_{ry} \cdot \cos\theta \quad (3)$$

That is, there exists a geometric relationship between the horizontal and vertical coordinates of the center point of the rear axis, as shown in Equation (5).

$$\dot{x}_r \cdot \sin\theta = \dot{y}_r \cdot \cos\theta \quad (4)$$

There is a geometric relationship between the velocity components of the vehicle forward circumference center point velocity v_f in the x and y axes, as shown in Equation (6).

$$\begin{cases} v_{fx} = v_f \cdot \cos(\theta + \varphi) \\ v_{fy} = v_f \cdot \sin(\theta + \varphi) \end{cases} \quad (5)$$

That is, the relationship between the horizontal and vertical coordinates of the center point of the front axis and its velocity can be expressed as

$$\begin{cases} \dot{x}_f = v_f \cdot \cos(\theta + \varphi) \\ \dot{y}_f = v_f \cdot \sin(\theta + \varphi) \end{cases} \quad (6)$$

Combining Equations (3), (5), and (7) yields the angular velocity of the vehicle heading as

$$\dot{\theta} = \frac{v_f \cdot \sin\varphi}{l} \quad (7)$$

Equations (7) and (8) represent the geometric relationship that exists between the velocity components of the center point of the rear axis at x and y axes, as shown in Equation (9).

$$\begin{cases} \dot{x}_r = v_f \cdot \cos\theta \cdot \cos\varphi \\ \dot{y}_r = v_f \cdot \sin\theta \cdot \cos\varphi \end{cases} \quad (8)$$

There is a geometric relationship between the front axle centroid velocity and the rear axle centroid velocity of the vehicle, as shown in Equation (10).

$$v_r = v_f \cdot \cos\varphi \quad (9)$$

Normally, the front wheels mainly control the direction of the vehicle movement, and the rear wheels provide power to determine the running speed of the vehicle; therefore, in the process of the low-speed motion of the vehicle parking, the overall motion speed of the vehicle is replaced by the speed of the center point of the rear axle of the vehicle, which is noted as v, and the center point of the rear perimeter of the vehicle is taken as the reference point of the vehicle motion; then, the equation of motion state of the center point of the rear axle of the vehicle can be expressed as

$$\begin{cases} \dot{x}_r = v \cdot \cos\theta \\ \dot{y}_r = v \cdot \sin\theta \\ \dot{\theta} = \dfrac{v \cdot \tan\varphi}{l} \end{cases} \tag{10}$$

2.2. Parking Lot Plan Model

The indoor parking lot of a shopping mall in Xuanwu District, Nanjing, is selected as the reference model for this study, as shown in Figure 2. There is one entrance and two exits in the parking lot. The parking spaces are numbered from 1 to 85 and arranged counterclockwise from the periphery to the inner periphery. The lane width is 5.5 m. The types of parking spaces include parallel, perpendicular, and angled parking spaces. The orange diagonal lines are the parking walls, and the white diagonal lines are the no-parking areas with other uses.

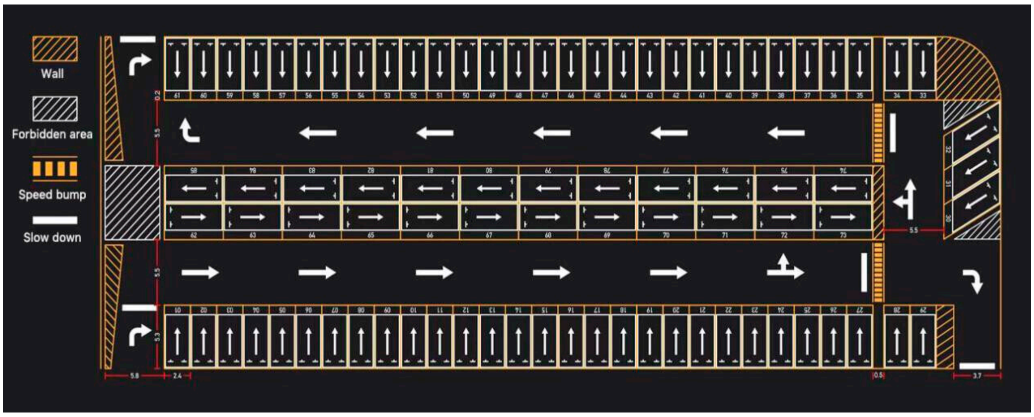


Figure 2. Parking lot plan.

2.3. Entropy Method [16–19]

The entropy weight method is an objective weighting method, the basic idea of which is to determine the objective weights according to the magnitude of the variability of the indicators. Based on the principle that the smaller the variability of an indicator, the less information it reflects, the lower the corresponding weight should be.

(1) Preprocessing of data: Assume that there are n objects to be evaluated, and m evaluation indicators (which have been normalized) constitute the normalization matrix as follows.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \tag{11}$$

The data are normalized, and the normalized matrix is denoted as Z. For each element in Z,

$$z_{ij} = x_{ij} / \sqrt{\sum_{i=1}^{n} x_{ij}^2} \tag{12}$$

This equation is used to determine whether there are negative numbers in the Z matrix, if so, another normalization method is needed for X. The matrix X is normalized once by the following formula.

$$\widetilde{Z_{ij}} = \frac{x_{ij} - \min\{x_{1j}, x_{2j}, \ldots, x_{nj}\}}{\max\{x_{1j}, x_{2j}, \ldots, x_{nj}\} - \min\{x_{1j}, x_{2j}, \ldots, x_{nj}\}} \tag{13}$$

(2) Calculate the weight of the ith sample under the jth indicator and consider it as the probability used in the relative entropy calculation; then, calculate the probability matrix P on the basis of the previous step, with each element in P as follows.

$$p_{ij} = \frac{\widetilde{z_{ij}}}{\sum_{i=1}^{n} z_{ij}} \tag{14}$$

(3) Calculate the information entropy of each indicator, calculate the information utility value, and normalize it to obtain the entropy weight of each indicator for the jth indicator, whose information entropy is calculated by the following formula.

$$e_{ij} = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln(p_{ij}) (j = 1, 2, \ldots, m) \tag{15}$$

The larger e_j is, the greater the information entropy of the jth indicator, and the smaller its corresponding information quantity. Defining the information utility value dj, the formula is as follows.

$$d_j = 1 - e_j \tag{16}$$

Normalize the information utility values to obtain the entropy weight of each indicator.

$$W_j = d_j / \sum_{j=1}^{m} d_j \tag{17}$$

2.4. Bayesian Networks [20–23]

Bayesian networks can be learned automatically directly from a database using empirically based algorithms that are usually built into appropriate software. Bayesian networks are well-suited for obtaining events that have occurred and predicting any one of several possible known causes.

Bayesian network model conditional dependence by representing conditional dependence through edges in the structure, and by extension, causation. Through these relationships and the use of factors, we can make effective inferences about random variables. The joint distribution of a Bayesian network is equal to the product of P (node | parent) of all nodes, as follows.

$$\begin{aligned} &P(X_1, \ldots, X_n) \\ &= \prod_{i=1}^{n} P(X_i | X_1, \ldots, X_{i-1}) \\ &= \prod_{i=1}^{n} P(X_i | Parents(X_i)) \end{aligned} \tag{18}$$

Inference in Bayesian networks takes two forms. The first approach is to simply evaluate the joint probability of a particular assignment of each variable (or subset) in the network. The second approach is to find P(x | e), or the probability of an assignment of a subset of variables given other variable assignments (x), where we must marginalize the joint probability distribution of variables that do not appear in x or e, which we denote as Y.

$$P(x|e) = \alpha \sum_{\forall y \in Y} P(x, e, Y) \tag{19}$$

2.5. Principle of KMO Test [24]

KMO is the abbreviation of Kaiser–Meyer–Olkin, which is an index to measure the close relationship between variables in factor analysis.

Suppose two samples are randomly selected from the overall sample pool $(X_i, Y_i)(i = 1, 2, 3 \cdots n)$; then, the coefficient of the simple linear relationship between the two samples is obtained as follows.

$$p = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}} \tag{20}$$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{21}$$

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{22}$$

The bias correlation coefficient is more able to reflect the intrinsic relationship between different variables by excluding the interference of other factors when the rest of the variables are determined. Equation (23) is the formula for the second-order bias correlation coefficient, and Equation (24) is the formula for the multi-order bias correlation coefficient.

$$r_{xy,z_1} = \frac{r_{xy} - r_{xz_1} r_{yz_1}}{\sqrt{(1 - r_{xz_1}{}^2)(1 - r_{yz_1}{}^2)}} \tag{23}$$

$$r_{xy,z_1,z_2\cdots z_h} = \frac{r_{xy,z_1,z_2\cdots z_{h-1}} - r_{xz_{h,z_1},z_2\cdots z_{h-1}}}{\sqrt{(1 - r_{xy,z_1,z_2\cdots z_{h-1}}{}^2)\left(1 - r_{yz_{h,z_1},z_2\cdots z_{h-1}}{}^2\right)}} \tag{24}$$

where r_{xy} is a simple correlation between X and Y, and $z_1, z_2 \cdots z_h$ is the identified h variable.

The KMO value of the sample data is expressed as the weight of the sum of the simple correlation coefficients of the sample variables in relation to the sum of the simple and partial correlation coefficients.

$$PP = \sum\sum_{i \neq j} r_{ij}{}^2 \tag{25}$$

$$RR = \sum\sum_{i \neq j} r_{ij,z_1,z_2\cdots z_k}{}^2 \tag{26}$$

$$KMO = \frac{PP}{PP + RR} \tag{27}$$

In 1974, Kaiser [25] published an article named 'An index of factorial simplicity' in the journal *Psychometrika*, which stipulated that the value of KMO was 0–1. For KMO $[0, 1]$, the closer the value to 1, the stronger the correlation between the variables, and the more suitable all the variables are for factor analysis; the closer the value to 0, the weaker the correlation between the variables, and the less suitable all the variables are for factor analysis. Table 1 shows the suitability of KMO for factor analysis at different values.

Table 1. Table of KMO test criteria.

Test Category	Range of Values	Factor Analysis is Appropriate for the Situation
KMO	$(0.9, 1]$	Perfect for
	$(0.8, 0.9]$	Great for
	$(0.7, 0.8]$	Suitable for
	$(0.6, 0.7]$	Barely fit
	$(0.5, 0.6]$	Not really suitable
	$[0, 0.5]$	Not suitable

3. Analysis and Results Discussion

3.1. Vehicle Parking Path Planning

For a visual demonstration, we chose parking spaces 10, 31, and 82 as examples, as shown in Figure 3.

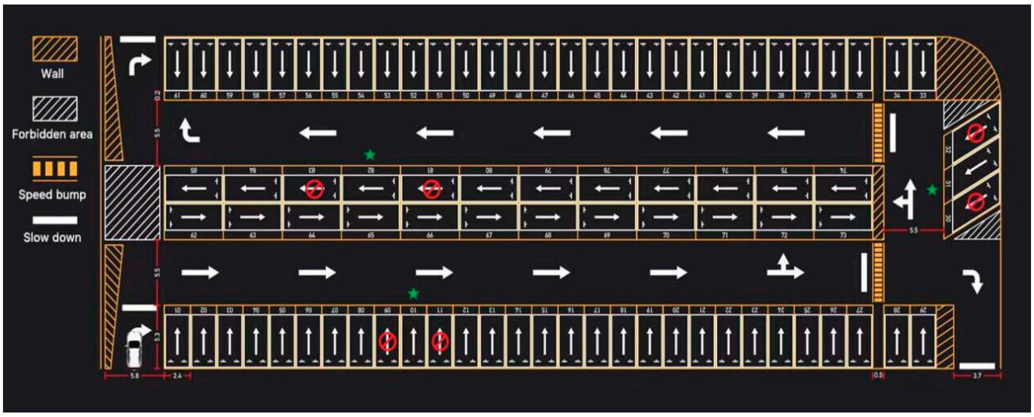


Figure 3. Selected vertical, horizontal, and diagonal parking positions.

3.1.1. Vertical Parking

Vertical Parking Path Planning

For the vehicle from the parking space entrance along a straight line, then through the arc trajectory into the horizontal section, and then forward a distance after the vehicle's control point to the right edge of the parking space with 11 in the same horizontal position. It began to reverse along the arc into the parking space and then along the straight line, it went forward a distance to the target parking position of such a reverse process using reverse thinking as shown in Figure 4, in order to facilitate the solution of the vehicle parking trajectory problem, simplify the complexity of the model, and give priority to the car in a straight line and arc on the reverse motion process; remember that the arc endpoints do not meet the endpoint of a straight line for the initial position.

When the initial position of the vehicle meets the requirement of one-step perpendicular parking, the parking-assisted parking system issues a command, and the vehicle owner can preferentially choose to reverse into the parking space according to the one-step perpendicular parking trajectory.

The geometric relationship between the turning radius of the vehicle R_1 and the horizontal coordinates of the initial position of the vehicle and the width of the parking space is reacted in Equations (28) and (29).

$$s_2 = s - s_1 \tag{28}$$

$$s_2 = R_1 \tag{29}$$

where s is the horizontal coordinate value of the initial position where the vehicle starts reversing into the $1/4$ arc.

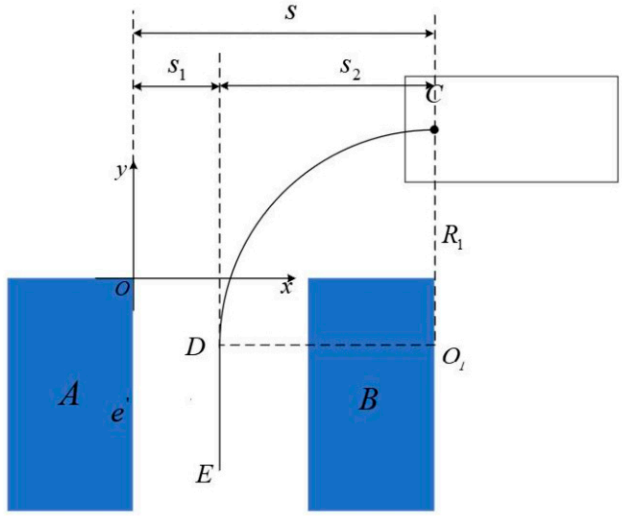


Figure 4. One-step vertical parking path planning diagram.

The relationship with the width of the parking space W_p is shown in Equation (30).

$$s_1 = W_p/2 \tag{30}$$

The relationship between the turning radius and the value of the horizontal coordinate of the initial position of the vehicle starting to reverse into the arc of 1/4 and the width of the parking space W_p can be obtained by combining Equations (28)–(30), as shown in Equation (31).

$$R_1 = s - W_p/2 \tag{31}$$

When the starting position of the vehicle does not meet the requirement of one-step vertical parking, a three-step vertical parking path planning is required, as shown in Figure 5.

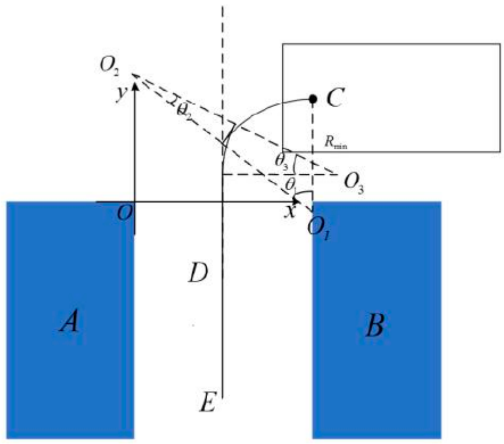


Figure 5. Three-step vertical parking path planning schematic.

For the three-step vertical parking path planning, given the limitation of parking space size and shape, the minimum turning radius is utilized as the radius of the circular arc trajectory when the vehicle reverses into the parking space, as shown in Figure 5. The vehicle quickly turns the steering wheel to the maximum turning angle from the starting position C at the maximum steering wheel speed to drive in reverse. When the control point of the vehicle, i.e., the center point of the rear axle, reaches the center axis of the 10th parking space, the vehicle no longer runs in reverse, but turns the steering wheel rapidly to the maximum turning angle at the maximum speed in the direction opposite to that of the backward direction for forward driving until the point E, and again turns the steering wheel rapidly to the maximum turning angle at the maximum speed to drive in reverse until the control point of the vehicle reaches the tangent to the center point of the rear axle. When the vehicle control point of the vehicle reaches the point F, which is tangent to the center point, turn the steering wheel to return to the original position, and the vehicle then moves in a straight line in the vertical direction along the center axis of the parking space 10 until it reaches the target parking point.

There is a numerical relationship between the three steering angles in Figure 5, as shown in Equation (32).

$$\theta_1 + \theta_2 + \theta_3 = \pi/2 \quad (32)$$

Analysis of Vertical Parking Obstacle Avoidance

For the vehicle parking process, multiple boundary conditions need to be considered as shown in Figure 6, such as the car cannot touch the boundary of the parking space, the road boundary, etc. In addition to meeting the basic conditions of not wanting to touch, the safety of the vehicle is fully considered, and the safety distance needs to be set for different boundaries. When the car is driven in the area delineated by these safety distance lines, the safety of the vehicle driving is improved, while the possibility of collision with the boundary is further reduced.

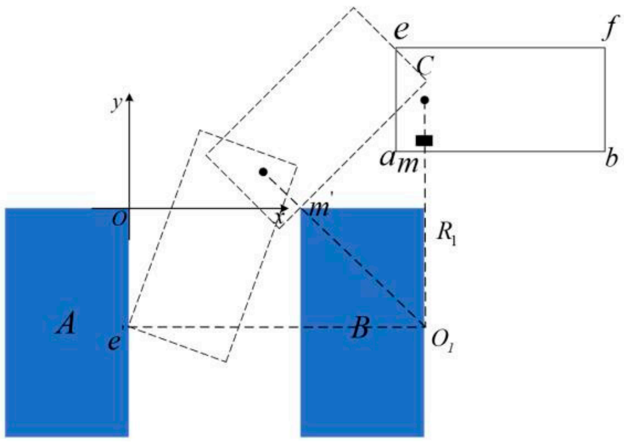


Figure 6. Analysis of the collision between the vehicle and the obstacles on both sides of the parking space.

Therefore, the following constraints need to be considered in the process of parking a vehicle into a parking area.

(1) The vehicle apex e does not collide with the left parking space boundary.

$$\begin{cases} x_{O_1} = O_1e' \\ O_1e' = O_1e \\ O_1e = \sqrt{(R_1 + w/2)^2 + {l_r}^2} \\ x_{O_1} \geq O_1e \end{cases} \quad (33)$$

(2) The vehicles do not collide with obstacles on the right.

$$\begin{cases} O_1P_1 = O_1m' \\ O_1m' = O_1m \\ O_1m = R_1 - w/2 \\ O_1P_1 = \sqrt{\left(x_{O_1} - W_p\right)^2 + (y_C - R_1)} \\ O_1P_1 \leq R_1 - w/2 \end{cases} \quad (34)$$

(3) The vehicles sit on top of the apex and road boundaries without colliding.

$$\begin{cases} O_1 f' = O_1 f \\ O_1 f = \sqrt{(R_1 + w/2)^2 + (l + l_r)^2} \\ O_1 f - R_1 + y_C \leq L_{road} \end{cases} \tag{35}$$

In order to reduce the possibility of the vehicle in the parking process and parking space edge or road edge and other obstacles' edge collision, in the analysis of the vehicle driving trajectory, the obstacles to add a safety distance d_3 are considered, and from the analysis of the minimum turning radius, parking space width, and other geometric factors, there is a certain value relationship, as shown in Equation (36).

$$\begin{aligned} \sqrt{\left(x_{O_1} - W_p\right)^2 + (R_{\min} - w/2 - d_3)^2 - \left(R_{\min} - W_p/2\right)^2} + d_3 \\ = R_{\min} - w/2 \end{aligned} \tag{36}$$

$$x_{O_1} = R_{\min} + W_p/2 \tag{37}$$

By calculating Equation (42), we can obtain the lower limit extreme value of the vertical coordinate y_c at the point C

$$y_C = R_{\min} - \sqrt{(R_{\min} - w/2 - \mathrm{d}_3)^2 - (R_{\min} - W_P/2)^2} \tag{38}$$

The horizontal and vertical coordinates of point D are as follows.

$$\begin{aligned} x_D &= W_p/2 \\ y_D &= -\sqrt{(R_{\min} - w/2 - d_3)^2 - \left(R_{\min} - W_p/2\right)^2} \end{aligned} \tag{39}$$

$$\begin{aligned} x_{C_{\max}} &= R_1 + W_p/2 \\ y_{C_{\max}} &= R_1 + y_D \end{aligned} \tag{40}$$

$$\sqrt{(R_1 + w/2)^2 + (l + l_r)^2} - R_1 + y_C \leq L_{road} \tag{41}$$

$$y_C = R_1 + y_D \tag{42}$$

$$x_{O_1} = \sqrt{(R_1 + w/2)^2 + {l_r}^2} \tag{43}$$

According to the spatial situation of the parking lot, the starting area where the vehicle may temporarily stay before parking is delineated using the parking space to be parked as the reference system, and four different boundary lines are determined based on the above boundary conditions for calculation, as shown in Figure 7. The expression of boundary line 1 is $(x - 2.4)^2 + (y - 4.98)^2 = 4.08^2$, the expression of boundary line 2 is $x = 5.973$, the expression of boundary line 3 is $y = 3.452$, and the expression of boundary line 4 is y = 3.01. When the vehicle is initially in the area below boundary line 4, the parking operation cannot be completed by one-step parking due to the limitation of parking space size, and the parking can only be completed by three-step parking. The vehicle can be parked in one step only when the vehicle control point position is located in the sector-like area enveloped by boundary line 1, boundary line 2, boundary line 3, and boundary line 4.

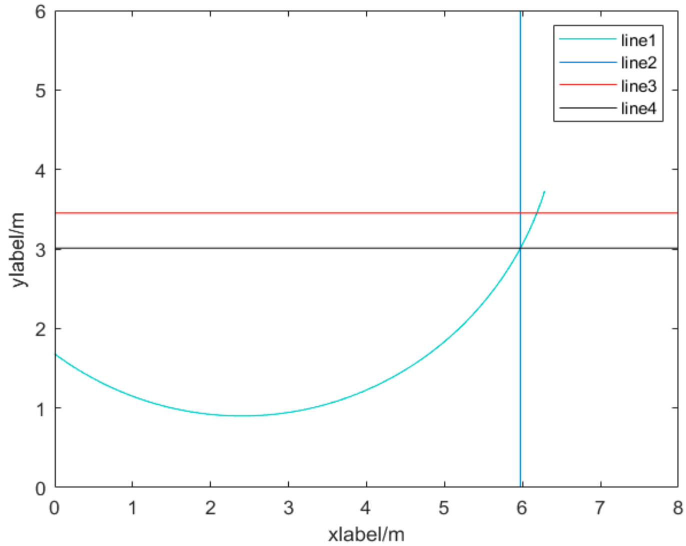


Figure 7. Vertical parking start area.

Vertical Parking Trajectory

In order to ensure safety and to make the time as small as possible, we choose one-step parking, i.e., the starting point of parking is located in the area below line 3 and above line 4.

The point in Figure 4 (5.6 m, 3.45 m) is selected as the starting point and modeled as follows.

t_1 is the time to do accelerated motion with changing acceleration, the velocity is v1, and the acceleration is a_1.

A uniform 1/4 circular motion in the time period t_2 is selected.

An accelerated motion with increasing acceleration in the time period t_3 is selected.

Uniform motion during the time period t_4 is selected.

t_5 is the time period for decelerating motion.

$$\begin{cases} \min t = t_1 + t_2 + t_3 + t_4 + t_5 \\ t_1 = v_1/a_1 \\ 0 \le a_1 \le 3 \\ -5 \le d_2 \le 0 \\ t_2 = \frac{1}{2} \cdot \frac{\pi R_k}{v_1} \\ \frac{da_1}{dt} \le 20 \\ v_2 = v_1 + a_1 t_3 \\ v_3 = v_2 + a_2 t_4 \\ v_3 = 0 \\ v_1 t_3 + \frac{1}{2} a_1 {t_3}^2 + v_2 t_4 + \frac{a_2}{2} {t_4}^2 = 2.4 \times 9 + s \\ t_5 = \frac{\theta R_k}{v_1} \end{cases} \tag{44}$$

By solving the above model, it can be seen that the vertical parking trajectory is shown in Figure 8.

The variation of speed, acceleration, and path of the vehicle from the beginning of the movement with time are shown in Table 2.

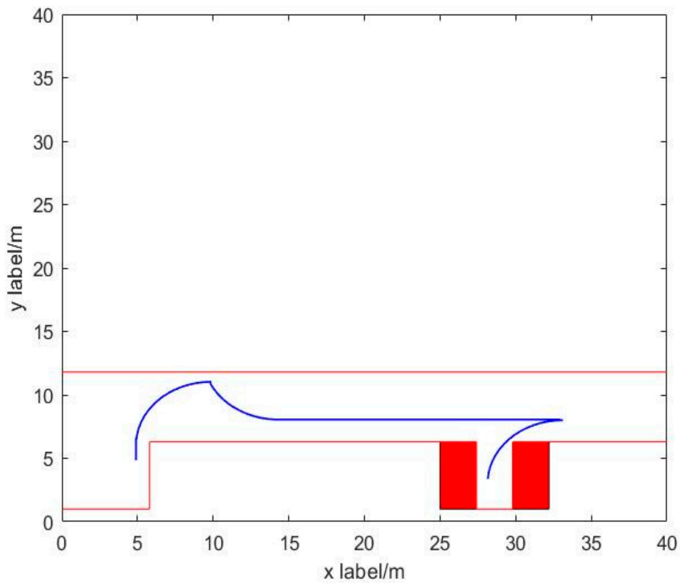


Figure 8. Vertical parking trajectory of the vehicle.

Table 2. Vehicle trajectory—time variation.

t/s	v/m/s	x/m	a/m/s^2	a'/m/s^3	θ/°
0	0	0	0	0	90
0.15	0.25	0.01125	3	20	90
0.59	1.544	0.4	3	0	90
7.082	1.544	10.423	0	0	0
7.232	1.769	10.655	3	20	0
8.494	5.556	20.055	3	0	0
10.619	5.556	31.863	0	0	0
10.919	4.656	33.348	−6	−20	0
11.695	0	35.155	−6	0	0

3.1.2. Parallel Parking

Parallel Parking Path Planning and Obstacle Avoidance Analysis

In order to make the parking path and time as short as possible, the turning radius of the last segment of the circular path of parking is set as the minimum turning radius, and the position of the starting point of the straight segment of the path is as close as possible to the origin of the coordinate system. First, the minimum parking space size for parallel parking is determined. As shown in Figure 9, the size of the minimum parking space required for parallel parking is mainly related to the second segment of the circular arc path, and the radius of the second segment of the circular arc path is the minimum turning radius to avoid collision as much as possible. Using the inverse planning method, it is assumed that the vehicle is parked at the target location of the parking space, and the vehicle can drive out of the parking space safely and smoothly according to the planned path FE and finally finish parking along the path. The vehicle does not collide with the P_1 point of the right obstacle in the process of driving out, then the minimum parking space length is calculated at this time as the minimum parking space length. Then, the vehicle b points to O_1 as the center of the circle, and R_b is the radius of driving to the b′ point. The b′ point and P_1 point distance should not be less than the safety distance d_1. According to the geometric relationship, we can derive the minimum parking space length L_{min} as follows.

$$L_{\min} = d_2 + l_r + \sqrt{R_b{}^2 - (R_{\min} - w/2)^2} + d_1 \tag{45}$$

$$R_b = \sqrt{(R_{\min} + w/2)^2 + (l + lr)^2} \tag{46}$$

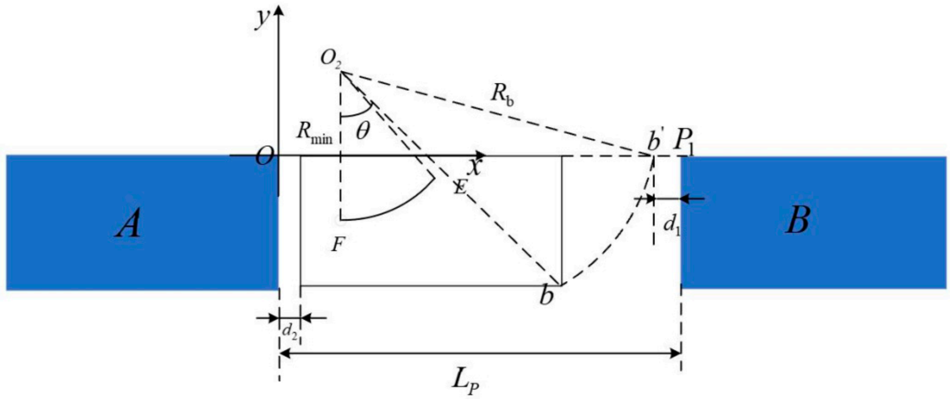


Figure 9. Minimum parallel parking space diagram.

When the safety distance d_1 tends to 0, the minimum parking space length L_{min} = 6.772 m, which is larger than the actual parking space length. Therefore, cars in the parking lot cannot be parked in one step but can only be parked in multi-step mode.

Parallel Parking Trajectory

According to the vehicle parking conditions and obstacle avoidance requirements, the following parallel parking model is established.

$$\begin{cases} \min T = t_1 + t_2 + t_3 + t_4 + t_5 + t_6 + t_7 \\ 0 \le a_1 \le 3 \\ -5 \le a_2 \le 0 \\ \dfrac{v_1{}^2}{2a_1} = 0.4 \\ t_2 = \dfrac{1}{2} \cdot \dfrac{\pi R_k}{v_1} \\ 0 \le \dfrac{da_1}{dt} \le 20 \\ -20 \le \dfrac{da_2}{dt} \le 0 \\ 0 \le \dfrac{da_3}{dt} \le 20 \\ -20 \le \dfrac{da_4}{dt} \le 0 \\ v_2 = v_1 + a_1 t_3 \\ v_3 = 2.77(10km/h) \\ t_4 = \dfrac{v_2 - v_3}{a_2} \\ v_1 t_3 + \dfrac{1}{2} a_1 t_3{}^2 + v_2 t_4 + \dfrac{1}{2} a_2 t_4{}^2 = 27 \times 2.4 - 5 \end{cases} \tag{47}$$

The model is solved to ensure safety and the smallest possible time, and the parallel parking trajectory is shown in Figure 10.

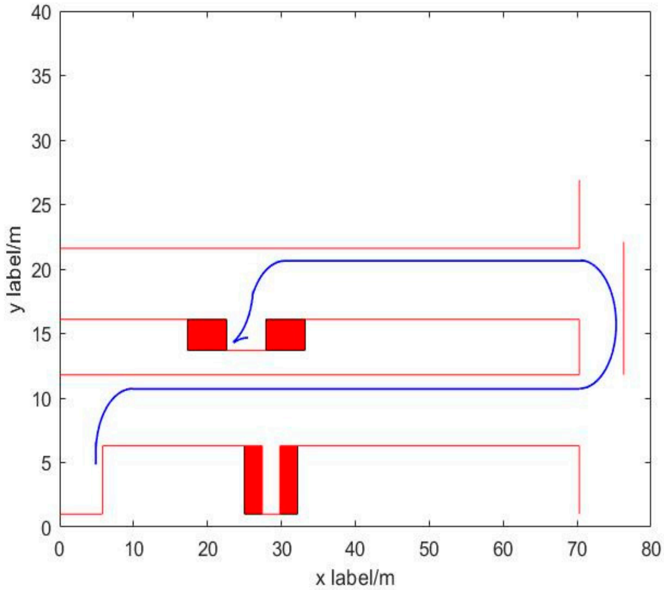


Figure 10. Parallel parking trajectory.

3.1.3. Inclined Parking

Inclined Parking Path Planning

As shown in Figure 11, the diagonal parking path consists of two parts: First, the target vehicle starts backing up in reverse at the starting point C, with the point O′ as the center of the circle and the turning radius R. When the vehicle heading angle is $\beta - \frac{\pi}{2}$ turn the steering wheel and start to back up along a straight line to the parking point E to complete the parking.

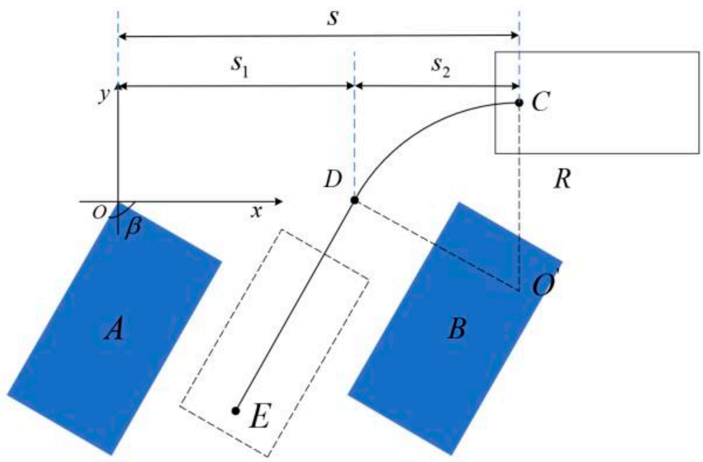


Figure 11. Oblique car parking path planning.

The turning radius is as follows.

$$R = \frac{s - s_1}{\sin \alpha} \quad (48)$$

where α is the turning angle during the vehicle movement, $\alpha = \pi - \beta$.

Obstacle Avoidance Analysis of Inclined Parking Path

When the target vehicle is being parked, two types of collisions may occur for parking situations where its final parking space axis position has been determined: the intersection position of the vehicle's rear axis extension line and the body may collide with the right obstacle, and the left front vertex of the vehicle may collide with the road boundary. The constraint conditions are as follows.

$$R_{\min} - w/2 \geq \sqrt{\left(x_c - x_{P_1}\right)^2 + \left(R_{\min} - y_c\right)^2} \tag{49}$$

$$L_{road} + R_1 - y_c \geq \sqrt{\left(R_1 + w/2\right)^2 + \left(l + l_f\right)^2} \tag{50}$$

Since the previous two parking methods were analyzed in detail and the analysis process of inclined parking is the same as shown in Figure 12, no specific analysis will be done here.

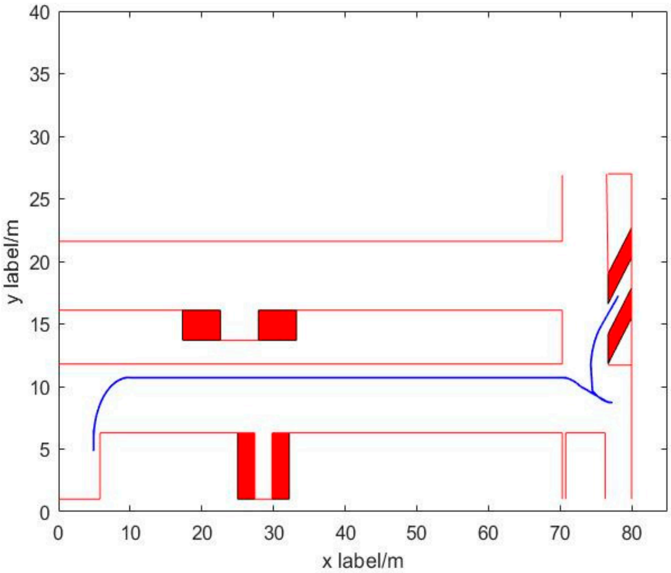


Figure 12. Inclined parking track.

3.2. Optimal Berth Selection

3.2.1. Static Optimal Parking Space Selection Based on Entropy Weight Method

Parking Space Attributes

The entire parking lot has 85 parking spaces as shown in Figure 13. Considering that spaces 1, 3, 4, 6–8, 10–12, 14–44, 45, 48, 50, 51, 55–63, 65, 66, 68–77, 79, 80, and 83–85, a total of 70 parking spaces, are red prohibited parking, meaning that they have been occupied, that leaves spaces 2, 5, 6, 13, 45, 47, 49, 52, 53, 54 64, 67, 78, 81, and 82, so a total of 15 free parking spaces are available.

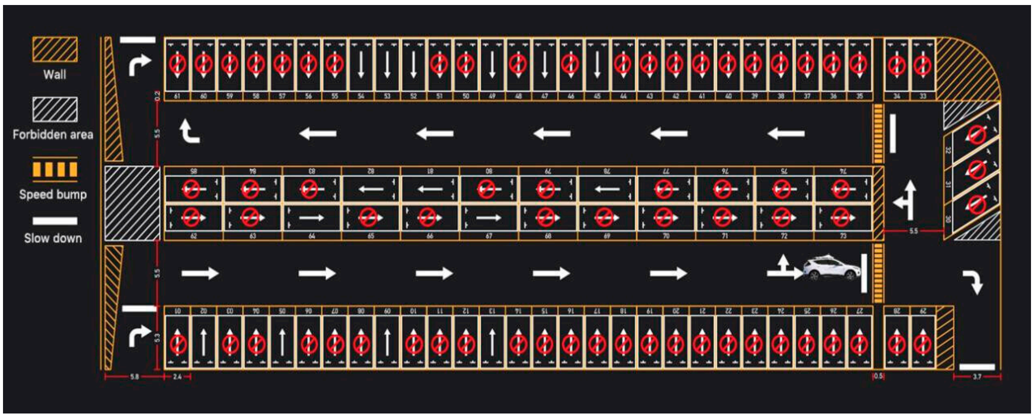


Figure 13. Parking space usage diagram.

Types of Parking Spaces

It is known that the vacant parking spaces are divided into three kinds of vertical parking spaces, parallel parking spaces, and inclined parking spaces. The difficulty of parking in these three kinds of spaces varies, from the perspective of practical experience, the difficulty of parking in vertical parking spaces, the difficulty of parking in parallel parking spaces, and the difficulty coefficient of parking in inclined parking spaces are all between [0 , 1], taking the following.

$$p_1 = \begin{bmatrix} \text{vertical parking difficulty,} \\ \text{parallel parking difficulty,} \\ \text{inclined parking difficulty} \end{bmatrix} \tag{51}$$

$$p_1 = [1, 0.5, 0.8] \tag{52}$$

At this point, the parking difficulty of the available parking spaces 2, 5, 6, 13, 45, 47, 49, 52, 53, 54, 64, 67, 78, 81, and 82 can be expressed according to their parking space type by the vector p_2 as follows.

$$p_2 = [0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1, 1, 1, 1, 1] \tag{53}$$

Left and Right Situation

From actual experience, it is known that the parking space left and right space occupancy also affects the parking difficulty. Generally speaking, the less the number of occupied parking spaces, the easier it is to park and the lower the parking difficulty factor; the more the target parking space is occupied, the more difficult it is to park and the higher the parking difficulty factor. The parking space is divided into four cases: left and right parking spaces are occupied, left parking spaces are occupied, right parking spaces are occupied, and left and right parking spaces are vacant. From the perspective of practical experience, the parking difficulty coefficients of these four situations are between [0, 1], and the orientation quantity p_1 indicates the parking difficulty coefficient of the above four situations for the left and right parking occupancy, as shown in the Equations (54) and (55):

$$p_3 = \begin{bmatrix} \text{left and right parking spaces occupied,} \\ \text{left parking spaces occupied,} \\ \text{right parking spaces occupied,} \\ \text{left and right parking spaces vacant} \end{bmatrix} \tag{54}$$

$$p_3 = [1, 0.6, 0.7, 0.5] \tag{55}$$

At this point, the parking difficulty of the available parking spaces 2, 5, 6, 13, 45, 47, 49, 52, 53, 54, 64, 67, 78, 81, and 82 can be expressed by the vector p_4 according to the occupancy of their left and right parking spaces as follows.

$$p_4 = [1, 1, 1, 1, 1, 1, 1, 0.6, 0.5, 0.7, 1, 1, 1, 0.6, 0.7] \tag{56}$$

3.2.2. Distance

In addition to considering the difficulty of parking with different types of parking and different occupied left and right parking spaces, the selection of the optimal parking space also needs to consider the distance from the starting point to the target parking space of the car and the walking distance that the occupant or driver needs to walk to leave the parking lot after parking is completed, so the distance from the starting point to different free parking spaces of the car needs to be included in the analysis.

In order to more conveniently represent the distance from the starting point of vehicle parking to the target parking space, a Cartesian coordinate system is established to conveniently represent the coordinate positions of different parking spaces and the distance from the starting point to the target parking space by combining the starting position with the location distribution of available free parking spaces. The coordinates of parking spaces are calculated by the center point of the parking spaces.

For this, the coordinates of the vehicle control point are set as the coordinate origin, the horizontal direction as the horizontal axis, the horizontal direction to the right as the horizontal axis positive, the vertical direction as the vertical axis, and the vertical direction up as the vertical axis positive.

The locations of the 15 available free parking spaces are represented by coordinates as shown in Table 3.

Table 3. Coordinates of vacant parking spaces.

Vacant Parking Space Number	Coordinate Location
2	(−57.35, −5.7)
5	(−50.15, −5.7)
6	(−40.55, −5.7)
13	(−30.95, −5.7)
45	(−21.35, 16.2)
47	(−26.15, 16.2)
49	(−30.95, 16.2)
52	(−38.15, 16.2)
53	(−40.55, 16.2)
54	(−42.95, 16.2)
64	(−44.6, 4.15)
67	(−29.3, 4.15)
78	(−19.1, 6.55)
81	(−34.4, 6.55)
82	(−39.5, 6.55)

Based on the coordinate locations of the free parking spaces in Table 3 and the direction of road access in the parking lot of this question, the distance between the location of these 15 available free spaces and the location of the starting point for the vehicle parking can be calculated as shown in the vector p_3, and the results are summarized in Tables 4 and 5.

$$p_3 = \begin{bmatrix} 57.63, 50.47, 40.95, 31.47, 26.80, \\ 30.76, 34.93, 41.45, 43.67, 45.90, \\ 44.79, 29.59, 20.19, 35.02, 40.04 \end{bmatrix} \tag{57}$$

Table 4. Parking attribute normalization matrix.

Parking Space Number	p1 (Type of Parking Space)	p2 (Left and Right Cases)	p3 (Distance)
2	0	1	1.00
5	0	1	0.81
9	0	1	0.55
13	0	1	0.30
45	0	1	0.18
47	0	1	0.28
49	0	1	0.39
52	0	0.2	0.57
53	0	0	0.63
54	0	0.4	0.69
64	1	1	0.66
67	1	1	0.25
78	1	0.2	0.00
81	1	0.4	0.40
82	1	1	0.53

Table 5. Entropy matrix of parking attributes.

Parking Space Number	p1 (Type of Parking Space)	p2 (Left and Right Cases)	p3 (Distance)
2	0.05	0.08	0.10
5	0.05	0.08	0.09
9	0.05	0.08	0.07
13	0.05	0.08	0.05
45	0.05	0.08	0.05
47	0.05	0.08	0.05
49	0.05	0.08	0.06
52	0.05	0.05	0.07
53	0.05	0.04	0.08
54	0.05	0.05	0.08
64	0.1	0.08	0.08
67	0.1	0.08	0.05
78	0.1	0.05	0.04
81	0.1	0.05	0.06
82	0.1	0.08	0.07

Combining the above-mentioned different types of parking methods and the difficulty of parking under the left and right occupancy of different parking spaces, the attribute information of parking in different vacant parking spaces is represented as matrix R_0:

$$R_0 = \begin{bmatrix} 0.5 & 1 & 57.63 \\ 0.5 & 1 & 50.47 \\ 0.5 & 1 & 40.95 \\ 0.5 & 1 & 31.47 \\ 0.5 & 1 & 26.80 \\ 0.5 & 1 & 30.76 \\ 0.5 & 1 & 34.93 \\ 0.5 & 0.6 & 41.45 \\ 0.5 & 0.5 & 43.67 \\ 0.5 & 0.7 & 45.90 \\ 1 & 1 & 44.79 \\ 1 & 1 & 29.59 \\ 1 & 0.6 & 20.19 \\ 1 & 0.7 & 35.02 \\ 1 & 1 & 40.04 \end{bmatrix}^T \quad (58)$$

Each column of matrix R_0 represents the attribute information of a parking space, including the type of parking space (vertical parking space, parallel parking space, inclined parking space), the left and right occupancy of the parking space (both left and right parking spaces are occupied, left parking space is occupied, right parking space is occupied, or both left and right parking spaces are free), and the distance between the location of the free parking space and the location of the starting point where the vehicle temporarily stays. Each row of the matrix R_0 represents one free parking space.

Based on the current parking space usage, the optimal parking space is selected from the remaining available free parking spaces. An evaluation matrix of the parking difficulty and driving walking distance of different parking spaces can be established, and the selection of the optimal parking space is determined by the evaluation scores of the different parking spaces. After the analysis, the comprehensive difficulty coefficient matrix of different available parking spaces R_0 is derived, and in order for it to be applied to the next factor analysis, it is first necessary to perform the *KMO* test.

After calculating the values of KMO for the evaluation matrix R_0, we get

$$KMO(R_0) = 0.4642 \tag{59}$$

According to the test table of KMO, it can be obtained that the correlation value between the variables of KMO is weak at less than 0.5, which is not suitable for factor analysis, so it is considered that there is no correlation between the attributes of each available vacant parking space, and there is no need to perform factor rotation and factor score calculation for R_0; instead, R_0 can be used directly as the original evaluation matrix of entropy weight method.

Table 6 shows the weights of the three attributes of parking space type, parking space left and right situation, and distance between starting point and parking space in the optimal parking space selection process calculated by the entropy weight method.

Table 6. Parking attribute weighting.

Properties	p1 (Type of Parking Space)	p2 (Left and Right Cases)	p3 (Distance)
Weights	0.517	0.211	0.272

The weight results of each attribute in Table 6 show that the weight of the parking space type is the largest, and the weight of the parking space left and right situation is the smallest, indicating that the parking space type has the greatest degree of influence on the final choice in the process of making the best parking space selection. Combined with the real situation, the type of parking space in the actual scenario largely affects the driving trajectory of the parking process and puts more space restrictions on the actual safe area where the vehicle can be parked compared to the parking space occupied by the left and right parking spaces. The walking distance of passengers and drivers is not very large because the parking lot is not very large, taking into account the limitation of its size and making the weight of the consideration of the walking distance not as heavy as the type of parking space in practice. Therefore, it is considered that the values of the weight of each attribute of the parking space obtained from Table 6 have a strong practical significance.

Table 7 shows the difficulty scores of each parking space after combining three factors: type of parking space, left and right use of parking space, and walking distance. The higher the difficulty score, the more difficult it is to park in this space, and the less likely it is to successfully park in this space and meet the walking distance requirement; conversely, the lower the difficulty score, the more difficult it is to park in this space.

Table 7. Parking difficulty score by parking space.

Parking Space Number	Difficulty Score
2	0.23
5	0.21
9	0.18
13	0.15
45	0.14
47	0.15
49	0.16
52	0.09
53	0.07
54	0.13
64	0.96
67	0.91
78	0.79
81	0.86
82	0.95

From Table 7, it can be seen that parking space 53 has the lowest overall parking difficulty score and the lowest difficulty factor and is the best parking space among the 15 free available spaces in this question.

When the user enters the parking lot, the parking-assisted parking system will inform the car owner that parking space 53 is the best parking space. Considering the actual situation that parking space 53 is a vertical parking space, the left and right parking spaces are not occupied and are free and available, there is a certain distance between the parking space position and the starting position—but it is not the farthest among the 15 free and available parking spaces, and the walking distance factor has the least weight among the three attributes, it is considered that parking space 53 is the best parking space.

3.2.3. Dynamic Optimal Parking Space Selection Based on Entropy Weight Method and Bayesian Network

As there are vehicles entering and leaving the parking lot every hour, the current parking space will be occupied or vacant at random. Although it is not possible to artificially determine the specific number of vehicles present in the parking lot and the parking space situation at the next moment, with the help of Bayesian networks, it is possible to account for and learn the probability of the available parking spaces and the number of parking spaces in the parking lot at the previous moment based on the computer clock (the optimal parking space is parking space 53 when there were only 15 free parking spaces among parking spaces 1–85 at the previous moment). This is used to infer the situation in the parking lot at the next moment. After the careful inference of Bayesian network, the new parking situation in the parking lot is obtained as shown in Figure 14, and then combined with the entropy weight method, so as to establish the dynamic optimal parking space selection model.

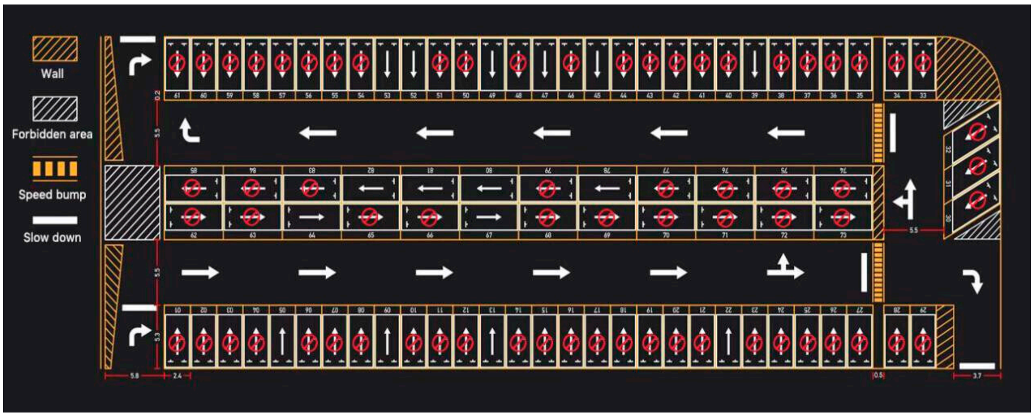


Figure 14. Bayesian network inferred parking space usage map for the new moment.

When the parking lot space usage changes dynamically, different space occupancy will affect the motion trajectory into the optimal parking space obtained from the above stationary state to some extent and change the parking difficulty while affecting the vehicle trajectory, so the already derived parking difficulty score table for each parking space needs to be updated.

First, the attribute information matrix R_0 for empty parking spaces is updated.

$$R_1 = \begin{bmatrix} 0.5 & 1 & 17.06 \\ 0.5 & 1 & 14.10 \\ 0.5 & 1 & 17.06 \\ 0.5 & 1 & 34.24 \\ 0.5 & 1 & 34.03 \\ 0.5 & 1 & 19.94 \\ 0.5 & 1 & 15.38 \\ 0.5 & 1 & 11.01 \\ 0.5 & 0.7 & 5.91 \\ 0.5 & 0.6 & 5.40 \\ 0.5 & 1 & 11.13 \\ 1 & 1 & 9.09 \\ 1 & 1 & 17.01 \\ 1 & 0.6 & 7.54 \\ 1 & 0.5 & 4.32 \\ 1 & 0.6 & 5.69 \end{bmatrix}^T \quad (60)$$

The updated R_0 will have the same meaning as the original R_0.

The updated matrix R_1 was first subjected to factor analysis, which after the *KMO* test yielded the following.

$$KMO(R_1) = 0.5271 \quad (61)$$

After observing the calculated values of *KMO* of the matrix R_1, we can find that the correlation of the matrix R_1 is very poor. It can be seen that the three factors selected here are independent of each other and have some reasonableness, so there is no need to do factor analysis, and R_1 can be used directly as the original evaluation matrix of the entropy weight method.

The weights of the three attributes of parking space type, parking space occupancy to the left and right of the parking space, and the distance between the target parking space and the starting position are reassigned by the entropy weight method, and the obtained results are shown in Figure 15.

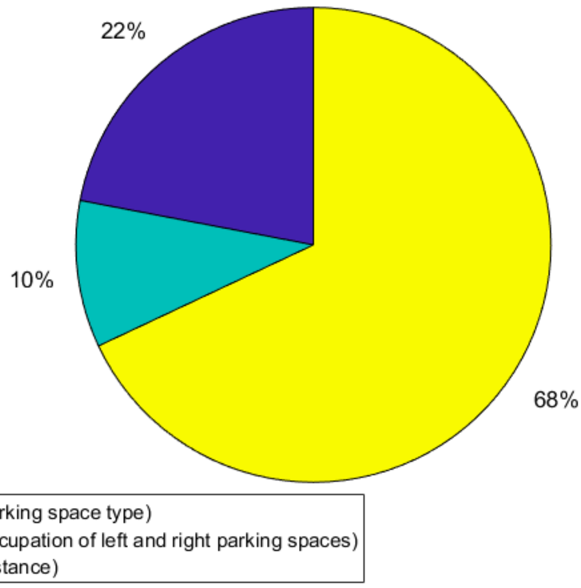


Figure 15. Updated parking space attribute weight chart.

After the update, the weight of the type of parking space is 0.22, the weight of the left and right occupancy of the parking space is 0.1, and the weight of the distance is 0.68.

The attribute with the highest weight value is distance. The distance of the vehicle from different parking spaces is plotted as shown in Figure 16, and this distance plays a dominant role in the score of the difficulty of parking all the vehicles.

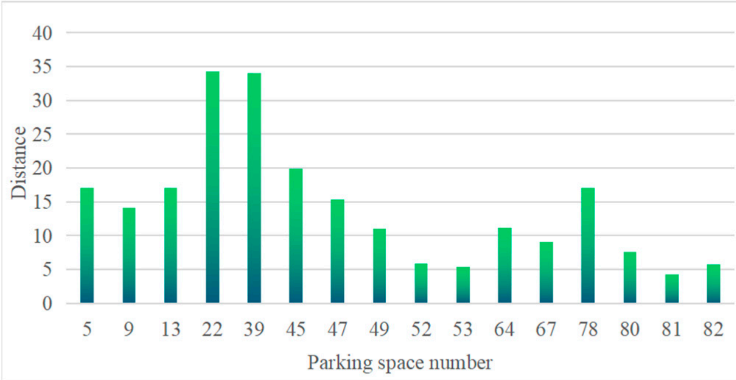


Figure 16. Updated distance map.

From Figure 17, the parking difficulty score shows that parking space 81 has the smallest comprehensive parking difficulty score and the lowest difficulty coefficient. Under the situation that vehicles enter and leave the parking lot every hour and each parking space will be occupied or released randomly, parking space 81 is the optimal parking space under the dynamic change of parking space occupation. The parking assistance system will indicate to the car owner that parking space 81 is the best parking space.

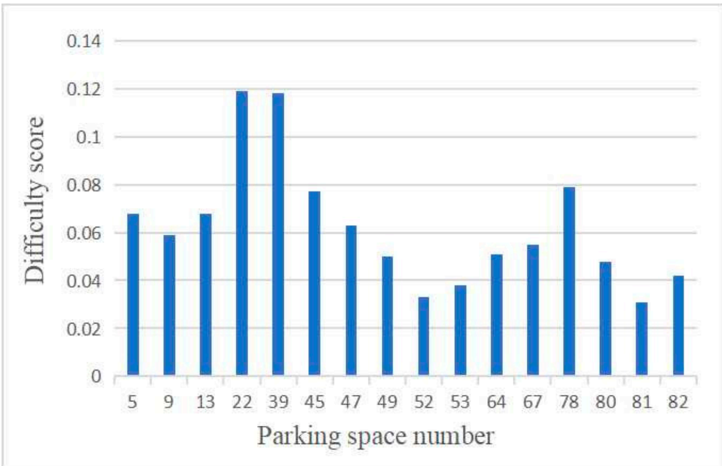


Figure 17. Updated difficulty score chart.

From the actual situation, parking space 81 is a parallel parking space, the left side of parking space 82 is not occupied—so it is in the free available state, the travel distance between parking space 81 and parking space 51 is close, and there are adjacent free parking spaces around, so parking space 81 is considered to be the best parking space in line with the actual situation and subjective understanding.

4. Conclusions

In summary, the proposed parking assistance method based on the entropy power method and Bayesian network for optimal parking and parking path planning is used for parking path optimization and selecting parking spaces. The main conclusions are summarized below:

(1) Clear and specific analysis of the obstacles when parking the vehicle into the parking space is demonstrated with good consistency, and the parking guidance trajectory given by the parking assistance parking system is both safe and fast.
(2) In the static case of selecting the optimal parking position, compared with other methods, the proposed method is more lightweight, faster, and more accurate in calculation, and can quickly give feedback on the optimal solution.
(3) In the dynamic case, the proposed method can learn the situation of the previous moment to reason out the parking space usage condition of the next moment, which is highly sensitive and can adapt well to the change of site conditions, and the auxiliary system could give the car driver a quick guide to find a suitable parking space in a crowded and busy parking lot.
(4) Based on the findings of this research, the efficient parking path and parking lot selection could be applied in the future rearrangement of indoor parking lots, which could be a feasible way to improve the whole process management sustainability.

Future works: The obstacle settings are all stationary, and no motion obstacle is considered. The speed of the parking process is certain, and the default of the whole parking process is a low-speed motion state, but if the vehicle driving speed is increased, the speed of the vehicle automatic obstacle avoidance turnaround needs to be optimized. At present, it is still in the theoretical design stage, and it is necessary to push the theoretical stage to the experimental stage and promote it to the public.

Author Contributions: Software, R.S.; formal analysis, J.Y. and D.G.; investigation, J.X. and J.Y.; resources, J.W.; data curation, J.X.; writing—original draft preparation, J.X. and R.S.; writing—review and editing, J.W., Y.W. and D.G.; supervision, Y.W.; project administration, K.H.; funding acquisition, K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by [University Student Practice Innovation Training Program Project-Nanjing Forestry University] grant number [2021NFUSPITP0215; 2021NFUSPITP0216; 2021NFUSPITP0205].

Data Availability Statement: All data can be asked from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, B.; Wei, Y.; Guo, I.Y. Automatic parking of self-driving car based on lidar. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-2/W7*, 241–246. [CrossRef]
2. Zhang, J.; Chen, H.; Song, S.; Hu, F. Reinforcement Learning-Based Motion Planning for Automatic Parking System. *IEEE Access* **2020**, *8*, 154485–154501. [CrossRef]
3. Gan, N.; Zhang, M.; Zhou, B.; Chai, T.; Wu, X.; Bian, Y. Spatio-temporal heuristic method: A trajectory planning for automatic parking considering obstacle behavior. *J. Intell. Connect. Veh.* **2022**, *5*, 177–187. [CrossRef]
4. Zhang, C.; Zhou, R.; Lei, L.; Yang, X. Research on Automatic Parking System Strategy. *World Electr. Veh. J.* **2021**, *12*, 200. [CrossRef]
5. Su, B.; Yang, J.; Li, L.; Wang, Y. Secondary parallel automatic parking of endpoint regionalization based on genetic algorithm. *Clust. Comput.* **2019**, *22*, 7515–7523. [CrossRef]
6. García, J.M.C.; Piorno, J.R.; Mata-García, M.G. Expert system design for vacant parking space location using automatic learning and artificial vision. *Multimed. Tools Appl.* **2022**, *81*, 38661–38683. [CrossRef] [PubMed]
7. Oetiker, M.B.; Baker, G.P.; Guzzella, L. A Navigation-Field-Based Semi-Autonomous Nonholonomic Vehicle-Parking Assistant. *IEEE Trans. Veh. Technol.* **2009**, *58*, 1106–1118. [CrossRef]
8. Zeng, C.; Ma, C.; Wang, K.; Cui, Z.; Dawson, K.A.; Indekeu, J.O.; Stanley, H.E.; Tsallis, C. Predicting vacant parking space availability: A DWT-Bi-LSTM model. *Phys. A Stat. Mech. Its Appl.* **2022**, *599*, 127498. [CrossRef]
9. Peng, X.; Sun, Z.; Chen, M.; Geng, Z. Arbitrary Configuration Stabilization Control for Nonholonomic Vehicle with Input Saturation: A c-Nonholonomic Trajectory Approach. *IEEE Trans. Ind. Electron.* **2021**, *69*, 1663–1672. [CrossRef]
10. Aye, Y.Y.; Watanabe, K.; Maeyama, S.; Nagai, I. An automatic parking system using an optimized image-based fuzzy controller by genetic algorithms. *Artif. Life Robot.* **2016**, *22*, 1–6. [CrossRef]
11. Zhang, P.; Yu, Z.; Xiong, L.; Zeng, D. Research on Parking Slot Tracking Algorithm Based on Fusion of Vision and Vehicle Chassis Information. *Int. J. Automot. Technol.* **2020**, *21*, 603–614. [CrossRef]
12. Chai, R.; Tsourdos, A.; Savvaris, A.; Chai, S.; Xia, Y.; Chen, C. Design and Implementation of Deep Neural Network-Based Control for Automatic Parking Maneuver Process. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *33*, 1400–1413. [CrossRef] [PubMed]

13. Nakrani, N.M.; Joshi, M.M. A human-like decision intelligence for obstacle avoidance in autonomous vehicle parking. *Appl. Intell.* **2021**, *52*, 3728–3747. [CrossRef]
14. Lee, C.K.; Lin, C.L.; Shiu, B.M. Autonomous Vehicle Parking Using Hybrid Artificial Intelligent Approach. *J. Intell. Robot. Syst.* **2009**, *56*, 319–343. [CrossRef]
15. Xie, J.; He, Z.; Zhu, Y. A DRL based cooperative approach for parking space allocation in an automated valet parking system. *Appl. Intell.* **2022**, *53*, 5368–5387. [CrossRef]
16. Zhang, H.; Jiang, W.; Deng, X. Data-driven multi-attribute decision-making by combining probability distributions based on compatibility and entropy. *Appl. Intell.* **2020**, *50*, 4081–4093. [CrossRef]
17. Joseph, A.G.; Bhatnagar, S. An online prediction algorithm for reinforcement learning with linear function approximation using cross entropy method. *Mach. Learn.* **2018**, *107*, 1385–1429. [CrossRef]
18. Jiaxin, L.; Weijun, W.; Yingchao, Z.; Song, C. Multi-Objective Optimal Design of Stand-Alone Hybrid Energy System Using Entropy Weight Method Based on HOMER. *Energies* **2017**, *10*, 1664.
19. Xu, T.; Gondra, I.; Chiu, D.K. A maximum partial entropy-based method for multiple-instance concept learning. *Appl. Intell.* **2016**, *46*, 865–875. [CrossRef]
20. Bgm, A.; Tdp, B. Advances in Bayesian network modelling: Integration of modelling technologies. *Environ. Model. Softw.* **2019**, *111*, 386–393.
21. Larra Aga, P.; Karshenas, H.; Bielza, C.; Santana, R. A review on evolutionary algorithms in Bayesian network learning and inference tasks. *Inf. Sci. Int. J.* **2013**, *233*, 109–125.
22. Rohmer, J. Uncertainties in conditional probability tables of discrete Bayesian Belief Networks: A comprehensive review. *Eng. Appl. Artif. Intell.* **2020**, *88*, 103384. [CrossRef]
23. Daly, R.; Shen, Q.; Aitken, S. Learning Bayesian networks: Approaches and issues. *Knowl. Eng. Rev.* **2011**, *26*, 99–157. [CrossRef]
24. Kim, J.O.; Mueller, C.W. Introduction to factor analysis: What it is and how to do it. *Contemp. Sociol.* **1980**, *9*, 562.
25. Kaiser, H.F. An Index of Factorial Simplicity. *Psychometrika* **1974**, *39*, 31–36. [CrossRef]

sustainability

Article

Two-Level Programming Model Based on Cooperative Operation Study of Stakeholders in Hazardous Chemical Storage

Jiao Yao *, Beibei Xie, Xiurong Wu and Cong Zhang

Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
* Correspondence: yaojiao@126.com

Abstract: Due to the uncertainty of risk occurrence and the severity of accident consequences in the process of hazardous chemical storage, there are many stakeholders involved in the management and supervision of hazardous chemical storage, and their interest appeals are different. On the basis of ensuring storage safety, in order to balance the interests of stakeholders and achieve cooperative operation, a two-level programming model considering the maximization of social welfare and the interests of warehousing enterprises was proposed. First, the upper model mainly refers to the regulatory department represented by the government, including the daily supervision cost, risk loss cost, risk compensation cost, and penalty coefficient formulated by combining various indicators. In the lower model, the comprehensive risk level of the warehouse is determined by the warehouse enterprise. Based on this, the supervision coefficient is determined. Combined with the punishment coefficient, the warehousing operation cost, warehousing supervision cost, and the punishment cost when the accident occurs under different risk levels are determined. The relevant case analysis shows that, compared with the evolutionary game model, the social supervision cost of the upper level and the enterprise cost of the lower level can be reduced by 0.49% and 30.43% respectively. Compared with the traditional improved particle swarm optimization algorithm, the proposed algorithm can reduce the supervision cost of the upper society and the lower enterprise by 0.11% and 7.05%, respectively, thus achieving a better supervision effect at a relatively low cost.

Keywords: hazardous chemical warehousing; two-level programming model; stakeholder; cooperative operation; penalty coefficient; improved adaptive particle algorithm

Citation: Yao, J.; Xie, B.; Wu, X.; Zhang, C. Two-Level Programming Model Based on Cooperative Operation Study of Stakeholders in Hazardous Chemical Storage. *Sustainability* **2023**, *15*, 1221. https://doi.org/10.3390/su15021221

Academic Editor: António Abreu

Received: 1 December 2022
Revised: 30 December 2022
Accepted: 5 January 2023
Published: 9 January 2023

1. Introduction

Due to the special physical and chemical properties of hazardous chemicals, special protective measures must be taken in the process of storage, transportation, production, and waste disposal. Otherwise, once an accident occurs, it will not only cause property loss, casualties, and environmental pollution, but also cause serious negative social impact. It is precisely because of the danger of hazardous chemicals that there are many stakeholders involved in the construction, operation, and supervision of hazardous chemicals storage facilities. Additionally, their interests are different, so there are multiple cooperative game problems. On the other hand, because the two-level programming model has the characteristics of hierarchy, independence, priority, and autonomy, it can effectively balance the interests of multiple stakeholders. Thus, the stakeholders can achieve the goal of cooperative operation under the premise of equality and mutual benefit of multiple stakeholders. This is also a kind of operation model that guarantees the normal operation of the main enterprises' rights and obligations.

Regarding the research on the risk of hazardous chemicals, scholars have adopted different research methods according to the characteristics of the subject. Tu Yuanyuan et al. used the cause theory of trajectory cross accidents to build a quantitative risk assessment index system in which the improved risk grading index method can more objectively reflect the risk of hazardous chemicals [1]. Zhang Le and Tong Xing analyzed the serious hazardous chemical accidents of 29 cranes in 2010–2019 so as to find a solution to promote safety production management [2]. Yue Baoqiang et al. established a power system risk assessment model for hazardous chemical explosion accidents in specific scenarios [3]. Yang Li et al. used electrostatic discharge to optimize the support vector machine so as to improve the dynamic risk assessment performance of the support vector machine model [4]. Ning Zhou et al. used the fuzzy comprehensive evaluation method to build a secondary evaluation index group of fire resource demand according to different storage methods of hazardous chemicals, which scientifically and effectively improved the fire fighting rate of hazardous chemicals [5]. Xiangcui Liu et al. designed a risk assessment and decision support system for hazardous chemicals transportation, mainly based on the needs of hazardous chemicals transportation risk assessment, path planning, and emergency rescue [6]. Wei Jiang et al., on the other hand, used the HFACS model to study the human–causal relationship of hazardous chemical storage accidents. It was concluded that the divergent claims of different interests in hazardous chemical storage are important causes of accidents [7].

Regarding the research on solving the divergent claims of interest subjects in hazardous chemical storage mostly adopts a game theory approach. Firstly, Wang Wei and Wang Xiaonan built a three-way evolutionary game model to study the impact of changes in fines, regulatory costs, and regulatory success rates on the evolutionary equilibrium, puting forward more targeted policy recommendations for multi-department cooperation [8]. Under the condition of asymmetric information, Xiao Maocai built a tripartite game model of regional government, local government, and hazardous chemical logistics enterprises in hazardous chemical logistics supervision to analyze the relationship of interests of the three in the process of hazardous chemical supervision [9]. Liu Jiaguo et al. introduced the incidence of safety accidents and studied the impact of port logistics enterprises and government regulatory authorities' strategy choices on the accident results [10]. Secondly, when the number of p-levels increases, the game strategy will change. Yingzuo Zhao et al. discussed the national supervision of hazardous chemical inspection institutions through evolutionary game on the basis of third-party chemical inspection [11]. Wang Wei et al. introduced public supervision into the supervision of hazardous chemical transport industry, proving that the degree of public participation is the decisive factor affecting the strategic choice of hazardous chemical transport enterprises [12]. It can be seen from the above literature that previous studies mostly affected the government's regulatory strategy through the formulation of game strategies by all parties, reducing regulatory costs and accident incidence. However, there is not an eternal zero-sum game relationship between the stakeholders, and the stakeholders can also coordinate the interests of all parties through a cooperative relationship.

Regarding the research on the promotion of collaborative stakeholder operating models, a review of the existing literature shows that two-level programming model is more widely used. Since the decision behaviors between the upper and lower layers in the two-level programming model are not completely constrained by each other, but they are adjusted in time according to the changes of the other side, one can finally making the model optimal as a whole, which may be applicable to the optimization of such problems. There are many studies on the two-level programming model to seek cooperation when the main objectives are different. For example, Mingtao Ma et al. used the two-level programming model to maximize the benefits of stakeholders sharing energy storage and reduce the operating costs of the hybrid system [13]. Long Yong et al. proposed a two-level programming model for the cooperative operation of relevant stakeholders of the micro grid, which can not only reduce carbon emissions, but also increase the profits of operators [14]. Zhou Ziyu and Jiang Huiyuan established a two-level programming model for cold chain logistics distribution location and path optimization, providing a theoretical basis for multi-objective optimization of the cold chain logistics network [15]. Zheng Bin et al. based on the characteristics of the post-earthquake disaster relief network, established a two-level programming dynamic model with the goal of maximizing the satisfaction of material delivery time at the upper level and maximizing the fairness of material distribution at the lower level [16]. Jianfeng Lu et al. established a two-level programming model for distribution and scheduling of emergency materials in hazardous chemicals storage to reduce the risk of hazardous chemicals storage in the emergency network [17]. Zhou Haixia et al. built a two-level programming model with the goal of minimizing the logistics cost at the upper level and the logistics time at the lower level so as to quickly and accurately transport emergency supplies to the demand point [18].

Therefore, the main contribution of this paper are as follows:

(1) This study focuses on the regulatory cooperation among the government, third-party regulatory agencies, the public, and hazardous chemical warehousing enterprises. It also analyzes the interest demands of each stakeholder.
(2) This study proposes a two-level planning economy model to balance the interest demands of all parties as much as possible.
(3) The model in this paper reduces social cost and enterprise cost to the greatest extent.

In general, game theory is generally used for interactive decision-making, where players choose strategies that benefit them. However, game theory is mostly non-cooperative games, and the two-level programming model makes up for it. Based on the above analysis, the social cost minimization pursued by the government, third-party regulatory agencies, and the public is combined with the overall cost minimization pursued by the hazardous chemicals warehousing enterprises. A two-level programming model, which is a linear programming problem, is established, and risk is introduced into the model in the form of cost [19]. The idea of the paper is divided into three steps as follows:

(1) Firstly, the government in the upper model formulates the relevant penalty coefficient, and the hazardous chemicals warehousing enterprises in the lower model formulates the supervision cost coefficient.
(2) Secondly, the hazardous chemicals warehousing enterprises in the lower-level model determine the comprehensive risk level of goods warehousing status according to relevant standards and estimate the probability of risk occurrence.
(3) Then, the penalty coefficient formulated by the upper-level government will affect the penalty cost of the lower-level enterprises. Through the formulation of the penalty coefficient, the lower-level warehousing enterprises are urged to meet the safety standards in daily supervision.

Based on the above analysis, the specific model structure is shown in Figure 1. Among them, in the upper level planning model, the following can be seen:

Daily supervision cost refers to the expenses incurred by the government and third-party regulatory agencies for the daily supervision of enterprises, which is to fulfill their own supervision responsibilities.

Risk loss cost refers to the cost that the government has to spend and the expected reduction of economic benefits, which is because of the existence of risks.

Risk compensation cost is the monetary compensation for the people who may bear the risk and suffer the injury around the hazardous chemical storage enterprise.

In the lower level planning model, it can be seen:

Warehousing cost refers to the sum of the input of various factors in the warehousing business activities of an enterprise and is expressed in the form of currency.

Warehousing supervision cost refers to the additional supervision cost invested according to the comprehensive risk level of different warehouses on the basis of daily supervision.

Estimated penalty cost is the penalty that the government imposes on the company after the accident according to the risk level of the accident.

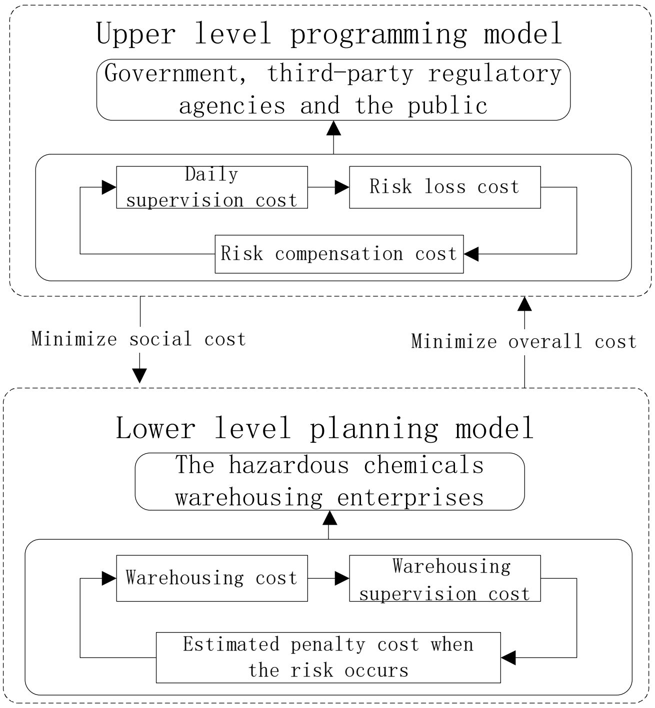


Figure 1. Model Structure of two-level programming model for stakeholders in hazardous chemicals storage.

2. Method and Model

This paper is divided into an upper and lower-level model. Relevant symbols are explained in Table 1. The penalty coefficient, α_{ri}, formulated by the government and the supervision cost coefficient, β_{ri}, formulated by the enterprise for different comprehensive risk levels of warehousing enterprises are taken as the decision variables to urge warehousing enterprises to reach safety standards in their daily management activities [20].

Table 1. Relevant symbols.

Symbols	Meanings	Ranges
α_{ri}	Penalty coefficient	$\alpha_{ri} > 0$
β_{ri}	Supervision cost coefficient	$\beta_{ri} > 0$
O_r	Risk compensation coefficient in warehouse r	$O_r > 0$
W_r	Storage of goods in warehouse r or not	$W_r = \begin{cases} 1, & \text{Storage of goods} \\ 0, & \text{No storage of goods} \end{cases}$
d_r	Comprehensive risk level	$\begin{cases} \textit{slight risk level}, & 0 < d_r \leq HD_r \\ \textit{medium risk level}, & HD_r < d_r \leq JD_r \\ \textit{serious risk level}, & d_r > JD_r \end{cases}$
R	Warehouses set	$R \in N+, r \in R$
Z_r	Daily supervision costs per unit area in warehouse r	$Z_r > 0, r \in R$
X_r	Storage area of goods in warehouse r	$X_r > 0, r \in R$

Table 1. *Cont.*

Symbols	Meanings	Ranges
U_r	Average risk loss in warehouse r	$U_r > 0, r \in R$
G_r	Risk compensation cost in warehouse r	$G_r > 0, r \in R$
U_{r1}	Loss when the estimated risk occurs in warehouse r	$U_{r1} > 0, r \in R$
U_{r2}	Fixed cost invested to deal with the risk in warehouse r	$U_{r2} > 0, r \in R$
E_r	Additional loss when the estimated risk occurs in warehouse r	$E_r > 0, r \in R$
F_r	Unit fixed cost invested in warehouse r to prevent the risk from occurring	$F_r > 0, r \in R$
ℓ_r	Additional unit loss incurred when the risk occurs in warehouse r	$\ell_r > 0, r \in R$
q_r	Estimated risk occurrence probability in warehouse r	$q_r \in [0, 1], r \in R$
N_r	Number of exposed populations within a radius of three kilometers	$N_r \in N+, r \in R$
b	Compensation for each person	$b \geq 0, r \in R$
t_r	Estimated risk value of warehouse r	$t_r \geq 0, r \in R$
$\bar{t}$	Average estimated risk value of warehousing enterprise	$\bar{t} \geq 0$
C_r	Unit storage cost of goods in warehouse r	$C_r > 0, r \in R$
P_r	Storage supervision cost in warehouse r	$P_r > 0, r \in R$
M_r	Estimated penalty cost when risk occurs in warehouse r	$M_r > 0, r \in R$
Y_r	Unit warehousing supervision cost in warehouse r	$Y_r > 0, r \in R$
L_r	Unit compensation cost	$L_r > 0, r \in R$
B_r	Maximum capacity of warehouse r	$B_r > 0, r \in R$
QL	Quantitative value of risk level	
D	Degree of risk impact	
Q	Probability of risk occurrence	$Q \geq 0$
RL_r	Comprehensive risk level in warehouse r	
ql_r	A quantitative value of risk level of a risk factor	
A_r	Risk weight value for a risk factor	

In view of the complexity of the research problem, the following assumptions are made for the model establishment:

(1) Considering the complexity of enterprise cost, this paper only studies the cost of enterprise storage, so as to replace the overall cost of the enterprise;
(2) The government involved in this paper is the general name of all government departments;
(3) This paper only considers one warehousing enterprise;
(4) In order to simplify the calculation, the classification of people is not considered in risk compensation;
(5) Assume that each costing is based on the unit area of hazardous chemical storage.

2.1. Construction of the Upper Level Model

The upper-level model mainly analyzes the daily supervision cost, risk loss cost, and risk compensation cost of the government, the third-party regulatory agencies, and the public in the process of participating in the supervision. The government estimates the risk value according to the comprehensive risk level so as to change the formulation of the penalty coefficient, as well as to affect the penalty cost of the lower level warehousing enterprises when the risk occurs. This is performed to ultimately reduce the probability of the final risk occurrence and, at the same time, minimize the social cost [21]. Specifically:

This is the upper social cost minimization objective function:

$$\min F = \sum_{r \in R} (Z_r X_r W_r + U_r + G_r) \tag{1}$$

where: Z_r is daily supervision costs per unit area of goods in warehouse, r, from the value of the government and the third-party regulatory agency; X_r is storage area of goods in warehouse, r; W_r is whether to store goods in warehouse, r; U_r is the average risk loss in warehouse, r, and G_r is the risk compensation cost in warehouse, r.

U_r is calculated as follows:

$$U_{r1} = E_r + F_r X_r W_r \tag{2}$$

$$E_r = \begin{cases} \alpha_{r1} \ell_r X_r W_r, & 0 < d_r \leq HD_r \\ \alpha_{r2} \ell_r X_r W_r, & HD_r < d_r \leq JD_r \\ \alpha_{r3} \ell_r X_r W_r, & d_r > JD_r \end{cases} \tag{3}$$

$$U_{r2} = F_r X_r W_r \tag{4}$$

$$U_r = q_r U_{r1} + (1 - q_r) U_{r2} \tag{5}$$

where, U_{r1} represents the loss when the estimated risk occurs in warehouse, r, E_r represents the additional loss when the estimated risk occurs in warehouse, r; F_r is the unit fixed cost invested in warehouse, r, to prevent the risk from occurring; $\alpha_{r1}, \alpha_{r2}, \alpha_{r3}$ is the penalty coefficients corresponding to the slight risk level $(0, HD_r]$, the medium risk level $(HD_r, JD_r]$, and the serious risk level (JD_r, ∞), which are the decision variables of the upper level model [22]. Among them, it is impossible to have a risk of 0 in the actual production activities. Therefore, the case of the comprehensive risk level $d_r = 0$ does not exist. ℓ_r indicates the additional unit loss incurred when the risk occurs in warehouse, r; U_{r2} represents the fixed cost invested to deal with the risk in warehouse, r; and q_r is used to describe the estimated risk occurrence probability in warehouse, r.

G_r is calculated as follows:

$$G_r = O_r N_r b \tag{6}$$

$$O_r = \left| t_r - \bar{t} \right| / \bar{t} \tag{7}$$

where O_r is the risk compensation coefficient in warehouse, r; t_r is the estimated risk value of the warehouse, r; $\bar{t}$ is the average estimated risk value of the warehousing enterprise;

N_r is the number of exposed population within a radius of three kilometers around the warehousing enterprises, and the number of surrounding population is obtained through GIS software; and b is the compensation for each person [23].

Constraint conditions:

$$W_r = \begin{cases} 1, & \text{Storage of goods in the warehouse } r \\ 0, & \text{No storage of goods in the warehouse } r \end{cases} \tag{8}$$

$$q_r \in [0,1] \tag{9}$$

$$\alpha_{ri} > 0, i = 1,2,3 \tag{10}$$

Equation (8) indicates the satisfaction state of the condition; Equation (9) indicates the probability range of risk occurrence; and Equation (10) indicates the value range of penalty coefficient.

2.2. Construction of the Lower Level Model

The lower-level model needs to ensure that the sum of storage cost, storage supervision cost, and penalty cost when the estimated risk occurs is minimized, that is, to minimize the overall cost of the hazardous chemicals warehousing enterprises. The objective function is as follows:

$$\min f = \sum_{r \in R} (C_r X_r W_r + P_r + q_r M_r) \tag{11}$$

where: C_r is unit storage cost of goods in warehouse, r; P_r is storage supervision cost in warehouse, r; M_r is the estimated penalty cost when the risk occurs in warehouse, r. P_r is determined according to the comprehensive risk level, d_r, and the enterprise supervision cost coefficient, β_{ri}; M_r is determined according to the comprehensive risk level, d_r, and the penalty coefficient α_{ri} of the corresponding level.

The storage supervision costs P_r is calculated as follows:

$$P_r = \begin{cases} \beta_{r1} Y_r X_r W_r, & 0 < d_r \le HD_r \\ \beta_{r2} Y_r X_r W_r, & HD_r < d_r \le JD_r \\ \beta_{r3} Y_r X_r W_r, & d_r > JD_r \end{cases} \tag{12}$$

where Y_r is the unit warehousing supervision cost in warehouse, r, and β_{r1}, β_{r2}, β_{r3} correspond to the enterprise supervision coefficients of slight, medium, and serious levels, respectively.

M_r is defined as (13):

$$M_r = \begin{cases} 0, & d_r = 0 \\ \alpha_{r1} L_r X_r W_r, & 0 < d_r \le HD_r \\ \alpha_{r2} L_r X_r W_r, & HD_r < d_r \le LD_r \\ \alpha_{r3} L_r X_r W_r, & d_r > LD_r \end{cases} \tag{13}$$

Among them, L_r is the unit compensation cost that the warehousing enterprise needs to pay when the estimated risk occurs in warehouse, r.

Constraint conditions:

$$\beta_{ri} > 0, i = 1,2,3 \tag{14}$$

$$W_r X_r \le B_r \tag{15}$$

Equation (14) indicates the value range of supervision coefficient; and Equation (15) indicates that the quantity of hazardous chemicals stored in a warehouse cannot exceed the maximum capacity of the warehouse, B_r.

3. Model Solving and Case Analysis

3.1. The Solution Method of Two-Level Programming Model

Based on the analysis of some existing excellent algorithm ideas [24], this study proposes to use the improved adaptive particle swarm optimization algorithm to solve the two-level programming problem. The parameters of the model were set as follows: population size $n = 50$, dimension $D = 2$, learning factor $c_1 = c_2 = 2$, and ωMax and ωMin are, respectively, 0.9 and 0.4. The termination condition of the algorithm is the maximum number of iterations, $iterMax = 50$, and the convergence accuracy $\lambda_\alpha \leq 10^{-10}$, $\lambda_\beta \leq 10^{-10}$, where the convergence accuracy is solved as in Equations (16) and (17).

$$\lambda_\alpha = \frac{|\alpha_{i+1} - \alpha_i|}{|\alpha_i|} \tag{16}$$

$$\lambda_\beta = \frac{|\beta_{i+1} - \beta_i|}{|\beta_i|} \tag{17}$$

Considering the randomness of the initial setting, the program was run independently for 30 times, and the best solution was taken as the approximate optimal solution. Additionally, the main improvement strategies of the algorithm are as follows:

(1) The disturbance factor is added to the velocity update formula to expand the population search range [25];
(2) The adaptive weight method is used to balance the global search ability and local improvement ability of PSO;
(3) Mutation operation is carried out on the global best solution, that is, random disturbance is added to the global optimum to improve the ability of PSO to jump out of the local best solution [26]. Assuming that random variable η follows the standard normal distribution, and its values are greater than 0 and less than 1, and the improved optimal position of particles is:

$$P_g^* = P_g(1 + \eta) \tag{18}$$

First of all, the upper-level decision maker formulates an appropriate accident penalty coefficient according to the comprehensive risk level of the warehouse storage state. The lower level takes this decision variable as a parameter, substitutes the result into the formulation of supervision cost, obtains the approximate optimal solution within the possible range, and feeds back the approximate optimal solution to the upper level. This cycle is repeated for many times, and the upper and lower levels of the two-level programming are optimized synchronously, and the approximate global optimal solution of the two-level programming model is finally obtained. The specific model solving process is shown in Figure 2 [27].

3.2. Case Analysis

This paper takes Shanghai Beifang Storage and Transportation Group as an example, uses the improved adaptive particle swarm optimization algorithm to solve, and gives the government's penalty coefficient α_i and the enterprise's supervision coefficient β_i for different risk levels in order to minimize the upper and lower objective functions, and then it analyzes the results.

3.2.1. Case Data Survey

Two warehouses of Beifang Group are selected as the research objects, namely, the dust sodium chlorate warehouse (warehouse 1) and the liquid hydrogen peroxide warehouse (warehouse 2), to find a relatively reasonable penalty coefficient and regulatory coefficient to minimize the social cost and the overall cost of the enterprise.

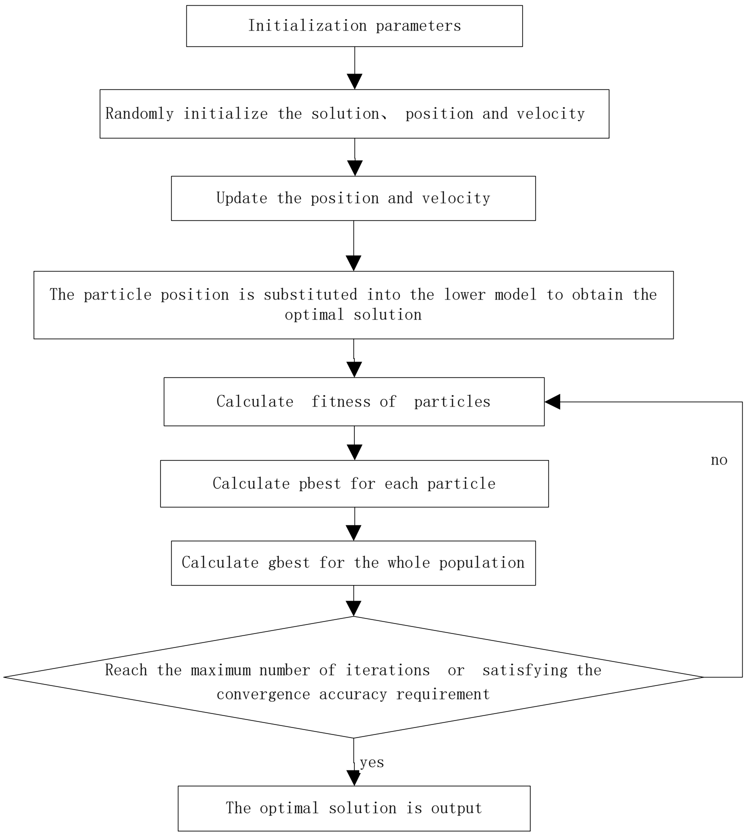


Figure 2. Solution process of two-level programming model.

In order to make the risk assessment results accurate and credible, expert survey method and the literature review were used to collect data and information. First, the quantified values of risk impact levels and the related criteria of risk impact degree were determined, as shown in Table 2. Then, based on the probability of risk occurrence and risk impact level, the risk level classification criteria and the level quantification interval were determined, as shown in Tables 3 and 4. From the relevant criteria in Tables 2–4, the quantitative value of risk level of each risk factor of hazardous chemical storage is calculated by using Equation (19).

Table 2. Risk occurrence probability and risk impact level table.

Risk Probability Level	Probability of Occurrence Q_1, Q_2	Influence Degree	Quantization Value of Influence Level D_1, D_2	Risk Impact Level
A	(0, 40]	Slight	(0, 4]	1
B	(40, 60]	Moderate	(4, 6]	2
C	(60, 100]	serious	(6, 10]	3

Table 3. Risk classification standard table.

Risk Probability Level	Risk Level		
	1	2	3
A	I	I	II
B	I	II	III
C	II	III	III

Table 4. Risk level quantification interval table.

Risk Level	Quantitative Range	Explain
I	(0, 1]	The risk is small and appropriate action is required
II	(1, 2]	The risks are high and prompt action is needed
III	(2, 3]	The risks are enormous and require immediate action

Among them, seven risk indicators affecting the safety of hazardous chemical storage were selected in a generalized manner by interviewing the relevant personnel of Beifang Group and combining with the General Rules for Storage of Commonly Used Chemical Hazardous Materials. Finally, the risk weights were calculated by applying the AHP rule to each risk factor in turn to determine the comprehensive risk level of each warehouse and complete Tables 5 and 6 [28].

Among them, since the quantization interval of the level is not a specific quantization value of the risk level, linear interpolation method is used for division, as shown in Equation (19):

$$QL = QL_1 + (QL_2 - QL_1)\frac{(D - D_1)(Q - Q_1)}{(D_2 - D_1)(Q_2 - Q_1)} \tag{19}$$

where: QL is the quantitative value of risk level, and QL_1, QL_2 are the value range of its quantization interval; D is the degree of risk impact, and D_1, D_2 are the value range of its quantitative value; Q is the probability of risk occurrence; and Q_1, Q_2 are the value range of its occurrence probability.

Table 5. Risk matrix table for warehouse 1.

Risk Factors	Probability of Risk Occurrence	Risk Impact		Risk Level		Risk Weighting A_r	Comprehensive Risk Level RL_1
		Quantitative Values	Level	Quantitative Values ql_1	Level		
Storage location is not reasonable	50	8	Serious	2.25	III	0.1429	II
Improper temperature and humidity control	90	9	Serious	2.57	III	0.2173	
Product deterioration	50	7	Serious	2.125	III	0.1825	
Mixed storage of chemicals of different properties	10	9	Serious	1.25	II	0.2021	
Lax control of ignition source	50	7	Medium	2.125	III	0.1002	
Lack of awareness of personnel management	40	7	Medium	1.25	II	0.1223	
Improper operation	30	3	Slight	0.57	I	0.0327	

In addition, the estimated risk occurrence probability of each warehouse can be determined by calculating the weighted average of each risk factor, and the calculation model is shown in Equation (20):

$$RL_r = ql_r A_r \tag{20}$$

where: RL_r is the comprehensive risk level in warehouse r; ql_r is a quantitative value of the risk level of a risk factor; and A_r is the risk weight value for a risk factor. Other relevant parameters in the investigation case are shown in Table 7, and the solution idea is shown in Figure 3.

Table 6. Risk matrix table for warehouse 2.

Risk Factors	Probability of Risk Occurrence	Risk Impact		Risk Level		Risk Weighting A_r	Comprehensive Risk Level RL_2
		Quantitative Values	Level	Quantitative Values ql_2	Level		
Storage location is not reasonable	20	1	Slight	0.125	I	0.0901	I
Improper temperature and humidity control	55	5	Medium	1.38	II	0.1437	
Product deterioration	50	5	Medium	1.25	II	0.1638	
Mixed storage of chemicals of different properties	10	9	Serious	1.25	II	0.2527	
Lax control of ignition source	30	3	Slight	1.57	II	0.0797	
Lack of awareness of personnel management	40	5	Medium	0.5	I	0.1522	
Improper operation	30	3	Slight	0.57	I	0.1187	

Table 7. Value of parameters.

Parameter	Symbol	Value	Unit
Daily supervision cost per unit of cargo	Z_r	12	RMB/ton
Storage capacity of hazardous chemicals in warehouse r	X_r	55 (Warehouse 1) 60 (Warehouse 2)	Ton
The additional unit loss incurred when the risk occurs	ℓ_r	120	RMB/ton
The fixed cost of dealing with the occurrence of a risk	F_r	50	RMB/ton
Compensation per person	b	1000	RMB
Unit storage cost of holding goods in warehouse r	C_r	10 (Warehouse 1) 9 (Warehouse 2)	RMB/ton
Unit storage supervision cost	Y_r	18	RMB/ton
The unit compensation to be paid when the risk is estimated to occur	L_r	300	RMB/ton

On the basis of Tables 5 and 6, the estimated risk probability of warehouse 1 is 48.73% and that of warehouse 2 is 32.43% by calculating the weighted average of all factors affecting warehouse risk. According to Equation (18), the estimated risk value of warehouse 1 is 1.904 and that of warehouse 2 is 0.998. Assuming that the average estimated risk value of warehouse 1 and warehouse 2 is taken as $\bar{t}$, the value is 1.451. Meanwhile, according to GIS, the number of exposed people within the radius of three kilometers of Beifang Logistics Company is 562. Ignoring the distance between warehouse 1 and warehouse 2, the risk compensation coefficient of warehouse 1 is 0.904, and that of warehouse 2 is 0.312.

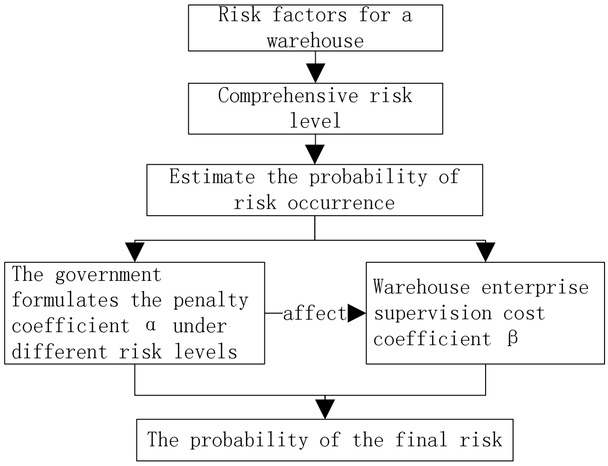


Figure 3. Case model solution ideas.

In addition, it is particularly emphasized that when the accident level is judged as serious, the enterprise itself is immediately rectified. Therefore, the case when the accident level is serious is not considered.

3.2.2. Case Result Analysis

MATLAB was used to program and solve the adaptive particle swarm optimization algorithm in Section 3. Considering the randomness of the initial setup, the program is run 30 times independently, and the best solution is taken as the approximate global optimal solution.

In addition, to further verify the effectiveness of the bilevel programming model for solving the cooperative operation research of the stakeholders of the hazardous chemical warehouse. Improved particle swarm optimization algorithm and classical evolutionary game method [29] are, respectively, used to compare the behaviors of regulatory entities represented by the government and warehousing enterprises in this paper. The specific numerical results are shown in Table 8, and the comparative analysis is shown in Figure 4. It can be seen that the optimal solutions of the adaptive particle swarm optimization algorithm in this paper are better than other solution methods. Not only the convergence performance of this algorithm is significantly improved, but also the standard deviation of the algorithm is 0 for 30 runs, and the results are very stable.

It can be seen that, in the case of a small degree of social cost optimization, the overall cost of the enterprise can still be significantly reduced. In addition, the method of this research can reduce the government's penalty coefficient for slight risk level to a large extent while ensuring the reduction of social costs and overall enterprise costs. However, it also increases the penalty for medium risk level. This is because an excessively high penalty factor may not produce a better supervision effect when the risk level is slight. On the contrary, it will bring great pressure to the enterprise operation and lead to unnecessary cost increases. However, when the risk level exceeds a certain range, a low penalty coefficient may also weaken the incentive for enterprises to supervise themselves, which increases the likelihood of risk occurrence. Therefore, a relatively better value should be obtained according to the actual situation so as to achieve the approximate optimal system configuration.

In addition, we can also see that in terms of enterprise supervision coefficient, the method in this paper can make enterprises strengthen their own supervision coefficient and fundamentally reduce the probability of accidents, no matter whether the risk level is slight or medium. At the same time, with the increase in risk level, the regulatory coefficient is also increasing. In conclusion, the results obtained by the cooperative operation model of the stakeholders of hazardous chemical warehousing based on two-level programming are more practical.

Table 8. Optimization results of model in this paper, evolutionary game model, and traditional particle swarm algorithm.

Plan	Social Cost/RMB	Overall Enterprise Cost /RMB	Slight Risk Level Government Penalty Coefficient (Warehouse 2)	Medium Risk Level Government Penalty Coefficient (Warehouse 1)	Slight Risk Level Enterprise Supervision Coefficient of (Warehouse 2)	Medium Risk Level Enterprise Supervision Coefficient of (Warehouse 1)
Evolutionary game model	698,073.7	28,756.8	2.2	1.47	1.49	1.41
Improper temperature and humidity control	55	5	Medium	1.38	II	0.1437
Traditional improved particle swarm optimization algorithm	695,398.9	21,524.1	1.05	1.47	0.76	1.66
Mixed storage of chemicals of different properties	10	9	Serious	1.25	II	0.2527
adaptive particle swarm optimization algorithm	694,612.9	20,006.7	0.24	1.82	1.31	1.51
Lack of awareness of personnel management	40	5	Medium	0.5	I	0.1522
Improper operation	30	3	Slight	0.57	I	0.1187

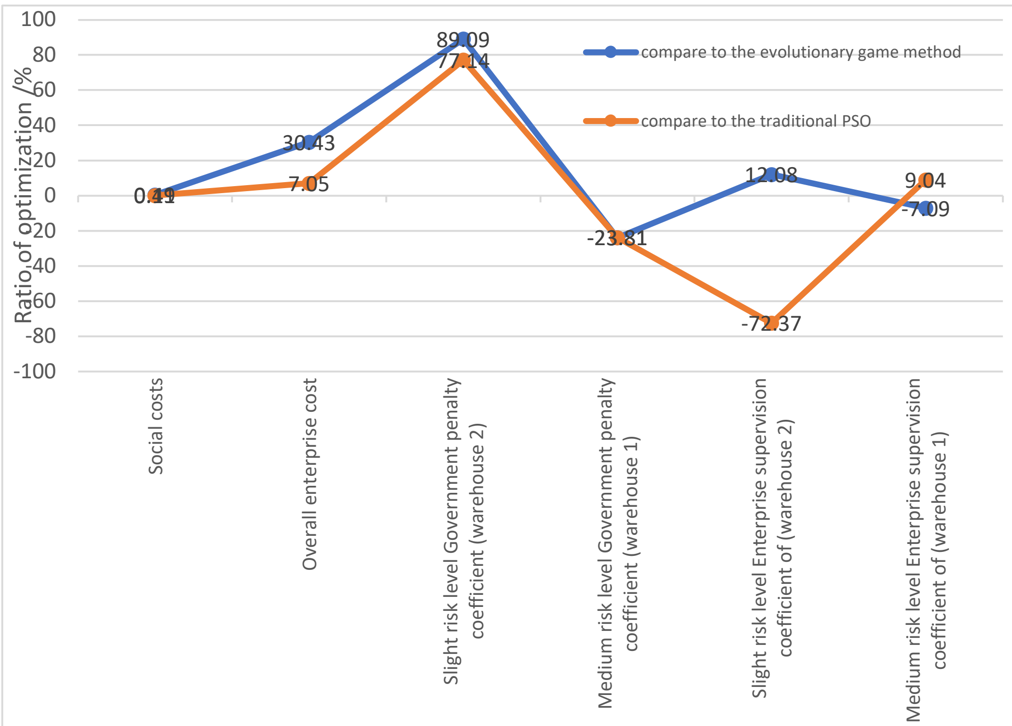


Figure 4. Comparison analysis of models among this paper, evolutionary game method, and traditional particle swarm algorithm.

4. Conclusions

Once an accident occurs in the hazardous chemical storage link, it will often cause catastrophic consequences. Moreover, there are many stakeholders, and it is difficult to balance the interests of each subject. Therefore, it is of great significance to study the way of cooperative operation among stakeholders. On the basis of summarizing the existing research results, this study analyses the cost composition of the government, third-party regulatory agencies, the public, and the hazardous chemicals warehousing enterprises in the process of production and operation under the premise that the hazardous chemicals warehousing enterprises may have an impact on society. This involves including the daily supervision cost and risk loss represented by the government and the warehousing operation cost, warehousing supervision cost, and the punishment cost when the risk occurs. A two-level programming model is established in which the upper level minimizes the social cost and the lower level minimizes the overall cost of the warehousing enterprise. Then, the improved adaptive particle swarm optimization algorithm is used to solve the model. Finally, the effectiveness of the model is verified by an example. The relevant case analysis shows that, compared with the evolutionary game model, the social supervision cost of the upper level and the enterprise cost of the lower level can be reduced by 0.49% and 30.43%, respectively. Compared with the traditional improved particle swarm optimization algorithm, the proposed algorithm can reduce the supervision cost of the upper society and the lower enterprise by 0.11% and 7.05%, respectively, thus achieving a better supervision effect at a relatively low cost.

However, only one company's data were validated by the model, so the validation of the same type of companies needs further improvement. Moreover, the relevant cost differences caused by hazardous chemicals with different chemical properties need to be discussed in the subsequent study. This can be performed by adjusting the composition of the cost parameters in the model and the costs incurred in the actual situation. Substituting these into the formula to solve realistic problems can improve the generalizability of the model. Additionally, through the model, government penalty coefficient and enterprise supervision coefficient for practitioners in the government or the enterprise were solved to develop a reasonable regulatory strategy to ensure the safety of hazardous chemical storage. Finally, about the proposed model in this paper, since two-level programming is an NP-hard problem, it needs to design a specific algorithm for the problem studied, so it does not have wide applicability.

Author Contributions: Conceptualization, J.Y.; methodology, B.X.; validation, X.W.; formal analysis, B.X.; investigation, C.Z.; resources, J.Y.; data curation, B.X.; writing—original draft preparation, B.X. and X.W.; writing—review and editing, C.Z.; visualization, B.X.; supervision, X.W.; project administration, C.Z.; funding acquisition, J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by MOE (Ministry of Education in China) Project of Humanities and Social Sciences (22YJAZH131) "Research on mechanism of route coordinated control of commuter traffic at the urban road network in the environment of big data". The authors, therefore, acknowledge with thanks the financial support of MOE (Ministry of Education in China).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tu, Y.Y.; Li, R.Q.; Wang, H.H.; Ciu, Y.; Liu, Y. Risk Assessment of Road Transportation of Hazardous Chemicals Based on Improved Grading Index Method. *J. Hunan Univ. Sci. Technol. Nat. Sci. Ed.* **2021**, *36*, 8–14. [CrossRef]
2. Zhang, L.; Tong, X. Prominent Contradictions and Solutions in the Safety Production Risk Governance: Analysis Based on 29 Hazardous Chemical Accidents. *J. Guangzhou Univ. Soc. Sci. Ed.* **2021**, *20*, 54–62. [CrossRef]
3. Yue, B.Q.; Sun, S.J.; Yang, L.C. Risk Assessment Model of Electricity System in Hazardous Chemicals Explosion Accident Based on Bayesian Networks. *J. Saf. Sci. Technol.* **2021**, *17*, 155–161. [CrossRef]
4. Li, Y.; Wang, H.; Bai, K.; Chen, S. Dynamic Intelligent Risk Assessment of Hazardous Chemical Warehouse Fire Based on Electrostatic Discharge Method and Improved Support Vector Machine. *J. Process Saf. Environ. Prot.* **2021**, *145*, 425–434. [CrossRef]

5. Zhou, N.; Xu, B.; Li, X.; Cui, R.; Liu, X.; Yuan, X.; Zhao, H. An Assessment Model of Fire Resources Demand for Storage of Hazardous Chemicals. *J. Process Saf. Prog.* **2020**, *39*, 1–8. [CrossRef]
6. Liu, X.; Huang, H.; Shi, H.; Zhang, T.; Zhao, J. Design of Risk Assessment and Decision Support System for Transportation of Hazardous Chemicals. *J. IOP Conf. Ser. Earth Environ. Sci.* **2021**, *692*, 1–7. [CrossRef]
7. Jiang, W.; Han, W.; Zhou, J.; Huang, Z. Analysis of Human Factors Relationship in Hazardous Chemical Storage Accidents. *J. Int. J. Environ. Res. Public Health* **2020**, *17*, 6217. [CrossRef]
8. Wang, W.; Wang, X.N. Research on Evolutionary Game of Hazardous Materials Transportation and Supervision Based on Multi-sectoral Collaboration. *J. China Saf. Sci. J.* **2020**, *30*, 69–74. [CrossRef]
9. Xiao, M.C. Research on Charging Strategy of Dangerous Goods Transportation under Multi-decision Subject. Master's Thesis, South China University of Technology, Guangdong, China, 2019. [CrossRef]
10. Liu, J.G.; Wang, J.J.; Zhou, H.; Zhang, Y.Q. Research on Supervision Problems of Port Hazardous Chemicals Base on Security Risk Level. *J. Syst. Eng.-Theory Pract.* **2018**, *38*, 1141–1152. [CrossRef]
11. Zhao, Y.Z.; Yuan, W.; Zhao, Z.; Li, W.; Li, H. National Supervision of Third Party Hazardous Chemical Inspection Institution by Evolutionary Game. *J. Chem. Eng. Trans.* **2021**, *71*, 1459–1464. [CrossRef]
12. Wang, W. Analysis on Evolution Game of Government and Hazardous Materials Transportation Enterprise Under Public Supervision. *J. Chin. J. Syst. Sci.* **2020**, *30*, 92–96.
13. Ma, M.T.; Huang, H.J.; Song, X.; Peña-Mora, F.; Zhang, Z.; Chen, J. Optimal Sizing and Operations of Shared Energy Storage Systems in Distribution Networks: A Bi-Level Programming Approach. *J. Appl. Energy* **2022**, *307*, 118170. [CrossRef]
14. Long, Y.; Liu, C.; Wang, Y. A Bi-level Programming Model Cooperatively Operated by the Microgrid Stakeholders. *J. Manag. Sci. China* **2019**, *22*, 69–82. [CrossRef]
15. Zhou, Z.Y.; Jiang, H.Y. Multi-objective Optimization of Cold Chain Logistics Network Based on Bi-level Programming Model. *J. Logist. Technol.* **2020**, *39*, 65–70. [CrossRef]
16. Zhen, B.; Ma, Z.J.; Zhou, Y.F. Bi-level Model for Dynamic Location-Transportation Problem for Post-earthquake Relief Distribution. *J. Syst. Manag.* **2017**, *26*, 326–337.
17. Lu, J.F.; Wang, X.X.; Zhao, J.H. Optimization of Emergency Supplies Scheduling for Hazardous Chemicals Storage Considering Risk. *Sustainability* **2021**, *13*, 10718. [CrossRef]
18. Zhou, H.X.; Mei, Y.R.; Lu, F.W. Research on the Optimization Method of Emergency Material Post Transportation Model Based on Bi-level Programming. *J. Chin. J. Internet Things* **2020**, *4*, 86–95.
19. Tan, Z.J. Capacity and Toll Choice of An Add-on Toll Road under Various Ownership Regimes. *J. Transp. Res. Part E* **2012**, *6*, 1080–1092. [CrossRef]
20. Wang, W.; Zhang, H.G. Transportation Network Optimization of the Hazardous Materials Considering Vehicle Speed Limit Interval. *J. Oper. Res. Manag. Sci.* **2021**, *30*, 128–134.
21. Hu, C.Y.; Liu, G.S. The Environmental Bi-Level Programming Models to the Location Decision in Distribution Centers. *J. Chin. Manag. Sci.* **2007**, *15*, 59–62. [CrossRef]
22. Liu, J.R. Open Vehicle Routing Optimization for Cold Chain Logistics Distribution Based on Bilevel Programming Model. *J. Bull. Sci. Technol.* **2019**, *35*, 209–212. [CrossRef]
23. Hu, J. Research on Solid Hazardous Waste Recovery Path Optimization considering Risk Equity. *Southwest Jiaotong Univ.* **2013**, *18*, 123–126. [CrossRef]
24. Song, Y.J.; Zhang, J.T. Multi-objective CuckooSearch Algorithm for Bi-level Programming Problems. *J. Oper. Res. Manag. Sci.* **2017**, *26*, 1–10.
25. Zhang, Y.R. Research on Solving Bilevel Programming Model Based on Improved Particle Swarm Optimization Algorithm. *Guangxi Univ.* **2013**, *18*, 123–126. [CrossRef]
26. Huang, W.Y. Optimal Bilevel Programming Model Based on Improved Particle Swarm Optimization Algorithm and Its Solution. *J. Stat. Decis.* **2018**, *1*, 88–91.
27. Zhao, Z.G.; Wang, W.Q.; Huang, S.Y. Bi-level Programming Problem Based on Improved Particle Swarm Algorithm. *J. Comput. Sci.* **2013**, *11*, 115–119. [CrossRef]
28. Ma, G.L. Study on Risk Assessment and Countermeasures of Hazardous Chemicals Storage in Petrochemical Enterprises. *J. China Econ. Trade Guide Theor. Ed.* **2020**, *10*, 128–130. [CrossRef]
29. Zhang, T.X. Analysis of Subject Behavior in Safe Production Supervision of Hazardous Chemicals Based on Game Theory. *Beijing Univ. Chem. Technol.* **2020**, *18*, 123–126. [CrossRef]

sustainability

Article

Study on Freeway Congestion Propagation in Foggy Environment Based on CA-SIR Model

Jiao Yao, Jiaping He, Yujie Bao, Jiayang Li and Yin Han *

Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
* Correspondence: hanyin2000@126.com

Abstract: The visibility in a foggy environment has a significant impact on driver behavior and traffic flow status, especially for whole closed highways with long distances between entrances and exits. Foggy days are very likely to cause congestion and even secondary traffic accidents, which seriously affect the reliability of freeway operation. In order to explore the influence of a fog environment on freeway traffic jams, firstly, this paper was based on the analysis of the impact of visibility on foggy days. Light fog, medium fog and heavy fog were classified as one scenario, while dense foggy weather was set separately as an extreme scenario without considering lane change. Furthermore, it used the SIR model of infectious disease for reference, and combined with the cellular automata (CA) model, the car-following model and lane-changing rules in different scenarios were set based on safe driving distance and speed for two scenarios. Finally, the key parameters of CA-SIR were calibrated, such as congestion propagation, recovery probability, vehicle braking, and lane-changing probability. The simulation analysis showed that with the decrease in visibility and vehicle speed, the phenomenon of congestion propagation was more prominent, but the causes of queuing phenomenon were different. A low speed limit was the main reason for traffic jams in the light fog condition. In the medium fog condition, the frequency of traffic jams was related to the random braking probability of the visibility. In heavy fog conditions, the congestion area gradually moved upstream with the passage of time. Moreover, in the dense fog condition, the congested area gradually moved upstream with the passage of time; however, vehicles were more likely to accompany each other, and the congested area traveled downstream synchronously with the passage of time and did not dissipate easily. Therefore, in a foggy environment, the best speed limit should be better established under different visibilities, the flow of highway traffic should be strictly controlled if necessary, and in worse situations than high-density traffic in low visibility, to avoid the spread of congestion, the intermittent release of different lanes is suggested to be implemented.

Keywords: traffic engineering; congestion propagation; foggy environment; cellular automata epidemic (CA-SIR) model; freeway

Citation: Yao, J.; He, J.; Bao, Y.; Li, J.; Han, Y. Study on Freeway Congestion Propagation in Foggy Environment Based on CA-SIR Model. *Sustainability* **2022**, *14*, 16246. https://doi.org/10.3390/su142316246

Academic Editors: Elżbieta Macioszek, Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 19 October 2022
Accepted: 2 December 2022
Published: 5 December 2022

1. Introduction

Adverse weather conditions such as fog, snow, heavy rain, and strong winds can negatively affect highways, especially in terms of road conditions, vehicle performance, visibility, and driver behavior. Drivers' driving behavior is sometimes difficult to quantify in response to such severe weather, and their perception is also affected by the impact of severe weather on their surroundings. The reduced visibility is more likely to increase the possibility of secondary accidents [1]. Wright and Roberg [2] found that severe weather is one of the fundamental causes of traffic congestion. Orosz et al. indicate that if a driver brakes or overtakes suddenly in certain limited situations, traffic congestion of up to 80 km may be observed, and the phenomenon is known as a "traffic tsunami" [3]. According to the Texas Transportation Institute [4], traffic congestion caused by episodic events such as severe weather accounted for 27% of the total number of traffic blockages. The report of the 2019 Blue Book of China Highway Network Operations showed that there were more than 20,000 traffic congestion blockages due to severe weather in 2019, and severe weather accounted for 17.4% of the causes of all congestion blockages. Among them, a foggy environment, as a more common type of severe weather, accounted for up to 9.5% [5], and the reduced visibility in a foggy environment leading to the restricted vision of drivers is an especially important factor affecting traffic flow characteristics [6]. Therefore, this paper attempts to carry out multi-scenario modeling and quantitative analysis of car following and lane changing to illustrate the effects of a foggy environment on the propagation of highway traffic congestion.

Firstly, a foggy environment has a significant effect on the behavioral characteristics of drivers. Zheng Shuxin et al. [7] studied the lane-change behavior of several drivers with different genders, ages, and driving ages through a high-simulation driving simulation cabin. It was found that as visibility decreases in a foggy environment, the lane-changing time increases while the lane-changing speed decreases, and as the reaction time before lane-changing increases, the car-following distance is shortened. Yan et al. [8] studied the effects of different risk levels on driver speed control in a foggy environment and found that, at a high risk level, the driver's slowed speed caused by the foggy environment was effective in reducing the severity of accidents, but it did not reduce the driver's risk of accident because drivers generally perceived that their own speed is slower than the actual speed in foggy weather. Zhang et al. [9] used Smart Eye to obtain information about the driver's gaze area, gaze angle, sweeping speed, and sweeping magnitude while driving in foggy weather and normalized the data to analyze the driver's gaze center of gravity. A series of conclusions were obtained: the driver's gaze speed fluctuated greatly when changing lanes on foggy days, and the gaze shift speed was significantly lower on foggy days than on sunny days.

Furthermore, regarding the effects of foggy conditions, extant research has focused on simulating driver behavior in foggy environments. There is a lack of research on driver behavior and performance in real-world environments. According to historical data from the Federal Highway Administration (FHWA) from 2007 to 2016 [10], about 15% of fatal crashes, 19% of crashes resulting in injuries, and 23% of crashes containing property damage occurred in severe weather. The challenge of low visibility, limited contrast, and perceptual distortion in foggy weather contributes to the large number of accidents that occur each year while driving in foggy weather. From a visual perspective, foggy environments can be described as involving a reduction in contrast in the vision field [11]. In a study of 566 interviewed drivers in Florida, Hassan et al. found [12] that when driving in foggy weather, drivers usually consider the vehicle in front as a means of guidance and maintain a similar speed to the vehicle in front. Particularly in conditions of limited visibility, drivers usually maintain a short headway from the vehicle in front. This tendency is thought to be the main reason for tailgating in foggy weather. Ahmed et al. [13] used various statistical methods to compare the behavior of drivers in foggy and sunny weather in order to better understand driver behavior in different weather conditions. The results showed that speed in dense fog and light fog was reduced by 10% and 2.8%, respectively, compared to the low speed in clear weather and that most drivers exceeded the speed limit in a given visibility condition.

In addition to the effects of foggy weather on drivers, macroscopic operating characteristics and patterns of traffic flow in this environment are also gradually receiving attention. Zhanhong Liu et al. [14] established a microscopic traffic flow model of foggy roads based on safety intervals using a cellular automata model. Additionally, car following in fog was simulated using the stochastic acceleration method, and the differences between fog and clear days were analyzed. Zhaowei Liu et al. [15] determined the effects of different levels of visibility on vehicle speed and the effect of road humidity on friction based on theoretical analysis for future unmanned driving in a foggy environment. Wenyan Feng et al. [16] chose the METANET traffic flow model to analyze the regional freeway network from a macroscopic perspective; moreover, they quantitatively analyzed the dynamic evolution of its driving conditions under different visibility levels. They also discussed the rule of influence of foggy days on the traffic operation of the regional freeway network so as to provide theoretical support for macroscopic traffic coordination and control of the regional freeway network.

The classic model for studying propagation is the SIR model. The SIR model and its modifications can be applied in different fields. Shah et al. used a generalized approach of SIR to accurately predict the spread of COVID-19-associated infections, recovery, and deaths in Pakistan [17]. The study by Miguel et al. showed that the SIR epidemic model can affect different communication layers of all nodes in a variety of Internet of Things (IoT) wireless networks [18]. According to the references, we can gather that the SIR epidemic model is more mature in the simulation studies of various types of propagation.

In summary, most of the existing studies have focused on the differences in driving behavior under foggy conditions and the effects on the operational characteristics of traffic flow. Regarding the congestion problem on the highway, prediction analysis is basically performed by the traffic flow model [19,20].

In contrast, there is a lack of systematic and in-depth investigations on traffic flow congestion, especially the evolution of the state of formation, propagation, and dissipation of freeway congestion. On the other hand, congestion propagation has many similarities with the classical SIR model of epidemics [21]. Each vehicle has three states: susceptibility (pre-crowding), affected (crowding), and exempted (post-crowding dissipation).

Therefore, the purpose of this paper is to study the congestion propagation mechanism of freeways more scientifically with a new method so that a theoretical basis for developing a control plan for freeway congestion in foggy environments can be provided. To better investigate the evolutionary details of propagation, this study combines it with the cellular automata (CA) model and divides the foggy environment links into two types of scenarios according to visibility.

Accordingly, the main contributions of this study are as follows:

(1) An analysis of the effect of visibility in a foggy environment;
(2) The establishment of the freeway congestion model based on the CA-SIR model in different foggy scenarios;
(3) The determination of key parameters of the CA-SIR model in a foggy environment;
(4) The case study and verification by MATLAB.

2. Model and Methods

2.1. Fundamental SIR Model

Lorenzo Pellis et al. verified that stochastic epidemic models can be extended to many other realistic models under specific conditions [22]. The model is able to ignore the temporal dynamics of the epidemic without affecting the final size distribution. The SIR infectious disease model is applied to the process of traffic congestion propagation in highways with three types of states: susceptible to congestion (S), in congestion (I), and removed from congestion (R) with the independent variable of time t, indicated as: $S(t)$, $I(t)$, and $R(t)$, respectively.

Vehicles in congestion affect the remaining uncongested vehicles in the surrounding area by random probability with λ as the average propagation rate. At the same time, the vehicles move away from the congested area with the average recovery rate of μ ($I \rightarrow R$ transition). The vehicles affected by the congested vehicles also become congested vehicles through the average propagation rate λ ($S \rightarrow I$ transition).

In the SIR propagation model, assume that the total number of vehicles within the whole process is N, which is kept constant. $S(t)$ is the free-flowing vehicle, $I(t)$ is the congested vehicle, and $R(t)$ is the departing vehicle, and then the differential equation between the three is:

$$\begin{cases} \frac{dS}{dt} = -\lambda I(t)\frac{S(t)}{N} \\ \frac{dI}{dt} = \lambda I(t)\frac{S(t)}{N} - \mu I(t) \\ \frac{dR}{dt} = \mu I(t) \end{cases} \tag{1}$$

Set separately:
The proportion of free-flowing vehicles to the total number of vehicles at time *t*:

$$s(t) = \frac{S(t)}{N}$$

The proportion of congested vehicles to the total number of vehicles at time *t*:

$$i(t) = \frac{I(t)}{N}$$

The proportion of departing vehicles to the total number of vehicles at time *t*:

$$r(t) = \frac{R(t)}{N}$$
$$s(t) + i(t) + r(t) = 1$$

Further, the nonlinear differential equation between the three is obtained as follows:

$$\begin{cases} \frac{ds}{dt} = \frac{d\left(\frac{S}{N}\right)}{dt} = \frac{1}{N} \cdot \left[-\lambda I(t)\frac{S(t)}{N}\right] = -\lambda i(t)s(t) \\ \frac{di}{dt} = \frac{d\left(\frac{I}{N}\right)}{dt} = \frac{1}{N} \cdot \left[\lambda I(t)\frac{S(t)}{N} - \mu I(t)\right] = \lambda i(t)s(t) - \mu i(t) \\ \frac{dr}{dt} = \frac{d\left(\frac{R}{N}\right)}{dt} = \frac{1}{N} \cdot [\mu I(t)] = \mu i(t) \end{cases} \quad (2)$$

2.2. Improved SIR Model

Since the original SIR model is a macroscopic static mathematical model based on ordinary differential equations, it may not be suitable for describing vehicles in congested conditions, especially those in constant motion. In addition, the differential equations in the SIR infectious disease model oversimplify the complex random behavior and are relatively complicated to solve computationally. In contrast, the cellular automata model simplifies the proof and solution challenges. It also allows for a better setting of complex stochastic behavior. Simulations can be performed with both time and space congestion propagation characteristics. Therefore, combining the two, an improved discrete model of CA-SIR is proposed.

Cellular automata is a discrete model with discrete time–space and state. It consists of four parts: cell, cell space *L*, cell neighbor *K*, and cell rule *F*. That is, $CA = (L, Sd, K, F)$, where *Sd* is the state set of the cell. Combined with the above description, the CA-SIR model can be used to accomplish the traffic congestion propagation problem by the following definitions.

(1) Cell space: A one-dimensional cell space containing *N* cells is established, and a cell in the one-dimensional cell space represents a vehicle in the network. The state of the next cell is determined by the state of the current cell and its neighbors.
(2) Cell state ensemble: let $Sd^t_{i,j}$ be the state of the cell in row *i* and column *j* at time *t*. Set $Sd^t_{i,j} = \{0, 1, 2\}$, where 0 represents the susceptible to congestion vehicles (*S*); 1 represents the vehicles in congestion (*I*); and 2 represents the vehicles not affected by congestion (*R*).
(3) Neighborhood rules: Moore-type neighbors.
(4) Evolution rule of cell:
    (1) When $Sd^t_{i,j} = 0$, if there are congested vehicles around the vehicle, each congested vehicle is influenced with probability λ. If the influence is successful, then $Sd^t_{i,j} = 1$; otherwise, $Sd^{t+1}_{i,j} = 0$;
    (2) When $Sd^t_{i,j} = 1$, the congested vehicle in the unit time step with the probability of b is transformed into a noncongested impact vehicle with probability. If the influence is successful, then $Sd^t_{i,j} = 2$; otherwise, $Sd^{t+1}_{i,j} = 1$;

(3) When $Sd_{i,j}^{t} = 1$ and $Sd_{i,j}^{t+1} = 2$, the vehicle leaves the congestion area and is no longer affected by the congestion.

In the congestion propagation process of freeways, the vehicle that generates congestion may potentially affect more subsequent vehicles. The more vehicles surrounding it, the faster the speed of congestion propagation and the wider the range of vehicles involved in the congestion.

2.3. CA-SIR Model of Freeway Congestion Propagation in Foggy Environment

The SIR infectious disease model is more mature in the simulation studies of various types of propagation. On the other hand, the cellular automata model generates propagation changes in the next moment through the interaction of individuals. In terms of traffic congestion propagation on freeways, while the assumptions of the SIR model cannot reproduce the actual traffic phenomenon, the cellular automata model can simulate the details of propagation through the evolution of its rules. The advantages of the two models combined can make the simulation closer to reality and more conducive to the analysis of the propagation process.

2.3.1. CA Model Setup for Highways and Visibility Effects

The classical one-way three-lane freeway is used as the model setting [23], where there are 500 cells within each lane, each cell represents one vehicle, and the cell length is set to 5 m so that the actual lane length is 2.5 km. There are two possibilities for each cell at any moment: occupied by vehicles or empty.

It is assumed that all vehicles in the model are small cars and that time, space, and vehicle speed are discretized by integers. The vehicles are randomly generated by Poisson distribution, the total number of vehicles N = 500, and from the beginning of the lane, the speed of the vehicle i at time t is $v_i(t)$ and $v_i(t) \in [0, v_{\max}]$, where $v_{\max}$ represents the maximum speed of the vehicle traveling. Since the set length of each cell is 5 m, it is established that $v_{\max}$ = 7 cell/s = 35 m/s = 126 km/h, which meets the speed limit requirement for driving on the freeway. Considering the influence of low visibility in different weather environments, the corresponding speed limit requirements in each visibility case in the cellular automata model were obtained, as shown in Table 1.

Table 1. The speed limit requirements corresponding to each visibility condition.

Definition	Visibility (m)	Speed Limit (km/h)	Speed Limit in the Cellular Automata Model (1 cell = 5 m)
Light	200–1000	80	20 (m/s) = 4 (cell/s)
Medium	100–200	60	15 (m/s) = 3 (cell/s)
Heavy	50–100	40	10 (m/s) = 2 (cell/s)
Dense	<50	20	5 (m/s) = 1 (cell/s)

Foggy weather can cause a significant change in visibility for drivers compared to sunny weather. This change also has an impact on traffic flow. While driving, drivers have 10–40% lower visual acuity in dynamic environments than in static environments. According to a related study [24], the relationship between meteorological visibility and a driver's visual distance can be obtained as shown in Equation (3):

$$L_s = \frac{0.6 d_q (\ln K + 3.912)}{3.912} \tag{3}$$

where L_s is the driver's visual distance, d_q is the visibility of the weather, and K is the contrast of the object itself. In the state of foggy weather, the object mostly appears gray or white, with the contrast generally taken as 0.35.

The driver's visual distance under foggy conditions is directly affected by visibility, and as the visibility range decreases, the driver's visual distance also decreases. In foggy weather environments, drivers cannot see the vehicle in front of them in the field of view, and at this time, the driving psychology is susceptible to change, and driving behavior tends to be conservative. The driver may adjust the speed within the visible distance in order to prevent the emergency braking of the vehicle when the distance ahead is small due to the low visibility.

2.3.2. CA-SIR Model for Different Fog Scenarios

(1) Scenario classification

According to the above study and the reasonable assignment for this study according to the relevant literature, the visibility distances were set to 400 m, 170 m, 75 m, and 40 m in light fog, medium fog, heavy fog, and dense fog, respectively [25], as shown in Table 2. The visibility distances corresponding to each visibility case are obtained from Equation (1). In this model, the cell length was set to 5 m, and therefore, the visible distance was taken down to a common multiple of 5, i.e., the driver's visible distances were 175 m, 75 m, 30 m, and 15 m. According to the minimum driving distance regulation of the Notice on Strengthening Traffic Management on Freeways under Low Visibility Meteorological Conditions of the Ministry of Public Security of China, the minimum safe distance d_{safe} for each visibility condition was finally obtained.

Table 2. Visibility distance and minimum safety distance corresponding to each visibility condition.

Definition	Visibility (m)	Vehicle Distance (m)	Visibility Distance L_s	Minimum Safety Distance d_{safe}
Light	400	>150	175 (m) = 35 (cell)	150 (m) = 30 (cell)
Medium	170	>100	75 (m) = 15 (cell)	75 (m) = 15 (cell)
Heavy	75	>50	30 (m) = 6 (cell)	30 (m) = 6 (cell)
Dense	40	—	15 (m) = 3 (cell)	15 (m) = 3 (cell)

On the freeway, the foggy environment, vehicle driving speed in different visibility conditions, and the safety distance that must be maintained are different, and the drivers' psychological characteristics and operational behavior are also different. Especially in extreme weather such as on dense fog days, the relevant implementation regulations provide that the maximum speed does not exceed 20 km/h and that drivers exit the freeway as soon as possible through the nearest exit to avoid secondary accidents. Due to the extreme weather, the driver's field of vision is extremely small, there is no reference system when driving on the freeway, and the lane-change situation is not considered; therefore, the basis for determining vehicle movement is also adjusted. In this study, the model was divided into two cases for research: light fog, medium fog, and heavy fog were grouped into scenario one; dense fog weather was set separately as scenario two.

(2) Vehicle following model of scenario 1

We integrated the corresponding visibility from Tables 1 and 2 to determine the speed limit conditions, vehicle distance, visual distance, and minimum safety distance under each visibility condition in the model. Different visibilities bring different sight distances, and driving speed has a significant impact on the driving state of the vehicle. In order to ensure the safety of vehicles on the highway and to avoid tailgating accidents in which the front and rear cars collide, a certain distance needs to be maintained so that the latter vehicle can have enough time to react, and this distance is called the safety distance.

According to Figure 1, there are 3 states of vehicles in the model proposed in this paper based on the different zones of front vehicles.

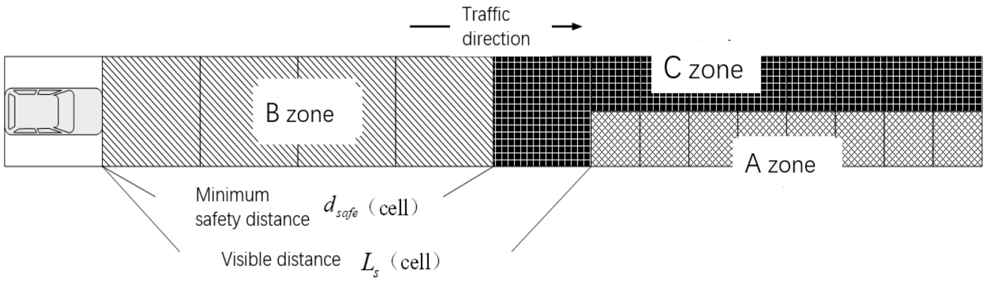


Figure 1. Schematic diagram of the area of the front car on a single lane of the freeway.

(1) The front vehicle is in the A zone (greater than the visible distance of current visibility). The rear vehicle can accelerate, and when the vehicle speed reaches the maximum speed, it can maintain the speed of driving until encountering the need to slow down.
(2) The front vehicle is in the B zone (less than the minimum safety distance of current visibility), and if the rear vehicle to continues to maintain its speed, it will cause rear-end collision with the front vehicle; therefore, the rear vehicle must slow down to maintain a safe distance.
(3) The front car is in the C zone (greater than the minimum safety distance of current visibility). The rear car makes a judgment according to the speed of the front car and the front car distance; therefore, there is a certain probability of a random deceleration behavior to adjust the speed.

Since a driver's effective sight distance is not necessarily within the safe distance, it is necessary to judge the distance of the car in front with the effective sight distance and the safe distance separately when building the model. To ensure the safety of driving in a foggy environment, drivers tend to be conservative in their behavior and adopt more deceleration behaviors and less acceleration behaviors.

The specific evolutionary rules are as follows.

(1) Judgment of the distance from the vehicle in front $d_{t,i}$ and the visibility distance L_s.

$$d_{t,i} = x_{t,i+1} - x_{t,i} - l_{car} \tag{4}$$

$$d_{t,i} = \begin{cases} L_s, & d_{t,i} \geq L_s \\ d_{t,i}, & d_{t,i} < L_s \end{cases} \tag{5}$$

where: $d_{t,i}$ indicates the distance between the first car and the car in front of it at the moment (cell); $x_{t,i}$,$x_{t,i+1}$ indicates the position of the car i and i + 1 at the moment (cell); l_{car} indicates the body length (cell), with l_{car}= 5 m; and L_s indicates the driver's visual distance.

(2) Vehicle state

1. Acceleration

A driver, when driving on foggy days, determines the acceleration of the vehicle according to whether there is a vehicle within the visible distance and whether the distance to the front car is greater than the minimum safe distance.

When there are vehicles in the same lane within the visual distance, the visual distance is greater than the distance in front, as shown in Figure 2a.

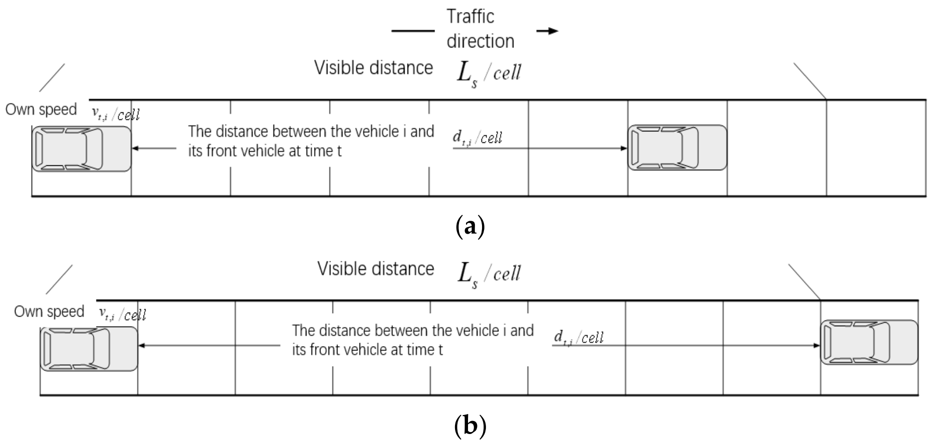


Figure 2. Visible distance and distance to the front vehicle on a single lane of the freeway. (**a**) Vehicles in the same lane within the visible distance; (**b**) No vehicle in the same lane within the visible distance.

When $d_{safe} < d_{t,i} \leq L_s$ and $v_{t,i} < d_{t,i}$,

$$v_{t,i} = \min(v_{t,i+1}, d_{t,i}, \alpha_n \cdot v_{\max}) \tag{6}$$

When there is no vehicle in the same lane within the visual distance, the visual distance is less than the distance in front of the vehicle, and the driver can accelerate within the visual distance, as shown in Figure 2b below.

When $d_{t,i} > L_s$ and $v_{t,i} < L_s$,

$$v_{t,i} = \min(v_{t,i+1}, L_s) \tag{7}$$

where: d_{safe} indicates the minimum safe distance (cell); $v_{t,i}$ indicates the speed of the first vehicle at the moment (cell/s); $v_{\max}$ indicates the maximum speed of the vehicle, i.e., the desired speed (cell/s); α indicates the speed coefficient under different visibilities; and $\alpha_n \cdot v_{\max}$ indicates the maximum desired speed (cell/s) under n visibility conditions. The smaller the visibility, the smaller the maximum expected speed; therefore, the specified value is within 0–1.

2. Random braking

A driver has the probability of performing small braking during nonacceleration and deceleration states.

$$v_{t,i} = \max(v_{t,i} - 1, 0) with P_{brake} \tag{8}$$

where P_{brake} is the random braking probability, which means that the vehicle has the probability P_{brake} of braking randomly.

3. Forced braking

When a driver is driving in fog and when there is a vehicle in the same lane within the visible distance, if the current vehicle speed at time *t* is larger or the distance from the vehicle in front is close to the minimum safe distance, the driver displays deceleration behavior:

$$d_{t,i} \leq L_s \text{ and } v_{t,i} > L_s, v_{t,i} = \min(v_{t,i} - 1, L_s) \tag{9}$$

4. Position update

$$x_{t+1,i} = x_{t,i} + v_{t,i} \times \Delta t \tag{10}$$

(3) Lane-changing rules for scenario 1 The number of vehicle lane changes in a foggy environment is significantly lower than that on a sunny day, and the lower the visibility, the smaller the probability of lane changes. When there is a slow vehicle

speed ahead or a vehicle in a congested state, and the speed of the adjacent lane is relatively fast, the subsequent vehicle may choose a suitable time to generate the lane-change behavior based on the visible distance, the vehicle speed, the following distance, and the speed of the neighboring vehicles in the adjacent lane. When the following conditions are met, vehicles may make a lane change.

(1) $d_{t,i} < \min\left(v_{t,i} + 1, d_{safe}\right)$. Indicates that the first vehicle is influenced by the vehicle ahead and probably will make a lane change.

(2) $d_{t,i,front} > \min\left(v_{t,i} + 1, d_{safe}\right)$. Indicates that there is enough lane-change space in the adjacent lane to provide a lane change for the first vehicle. $d_{t,i,front}$ is the distance (cell) between the first car and the nearest preceding car in the adjacent lane at the time.

(3) $d_{t,i,back} > \min\left(v_{t,i,back} + 1, \alpha_n \cdot v_{\max,i,back}, d_{safe}\right)$. $d_{t,i,back}$ is the distance between car i and the nearest car in the adjacent lane at time t (cell); $v_{t,i,back}$ is the speed of the nearest car in the adjacent lane at the time (cell/s); and $\alpha_n \cdot v_{\max,i,back}$ is the maximum speed of the nearest car in the adjacent lane at the time (cell/s) at visibility n.

(4) $rand() < \left(1 - P_{change}\right)$. P_{change} is the lane-change probability of the vehicle; $rand()$ is a random number between 0 and 1.

In this study, the lane change probabilities of vehicles in foggy environments were 0.41 (light fog), 0.23 (medium fog), and 0.14 (heavy fog).

(4) Vehicle following model of scenario 2

In a foggy environment, vehicles change lanes less frequently than in clear weather because they have lower visibility and shorter visual distance, and the judgment basis for lane change needs to be more precise than in clear weather.

Related research [26] shows that since freeways are generally in open areas, vehicles traveling in the same lane subconsciously shorten the distance to the vehicle in front to obtain the longest visible distance at that visibility in a foggy condition with 48 m visibility. The faster the vehicles in the adjacent lane, the faster the driver unconsciously maintains the same fast speed. The driver focuses on the vehicle in front, and it is difficult to observe the location of the rear vehicle through the rearview mirror. Therefore, for this kind of extreme weather environment (visibility of 50 m), ignoring the driver's individual lane change and considering the influence of vehicles in the adjacent lane, only the vehicle-following model under a dense fog environment was established.

In the model, according to the influence of the following distance and driving speed of the vehicle and the adjacent-lane vehicles, two parameters, γ and δ, are used to indicate the influence of the adjacent-lane vehicles on the following distance and speed of the vehicle, respectively. The larger γ is, the greater the impact of the distance difference between the current vehicle and the adjacent-lane vehicle on the driver in the visible range; the larger δ is, the greater the impact of the speed difference between the current vehicle and the adjacent-lane vehicle on the driver in the visible range. Based on the results of related studies [18], the values of 0.4 and 0.5 were taken for the two parameters, respectively.

Under consideration of the above conditions, in order to measure the influence of the left and right adjacent lanes on the braking probability of the vehicle, three unidirectional lanes were used as an example to study the congestion propagation phenomenon under a dense fog environment. The simulation diagram in a dense fog environment is shown in Figure 3. The following rules were set for the following operations.

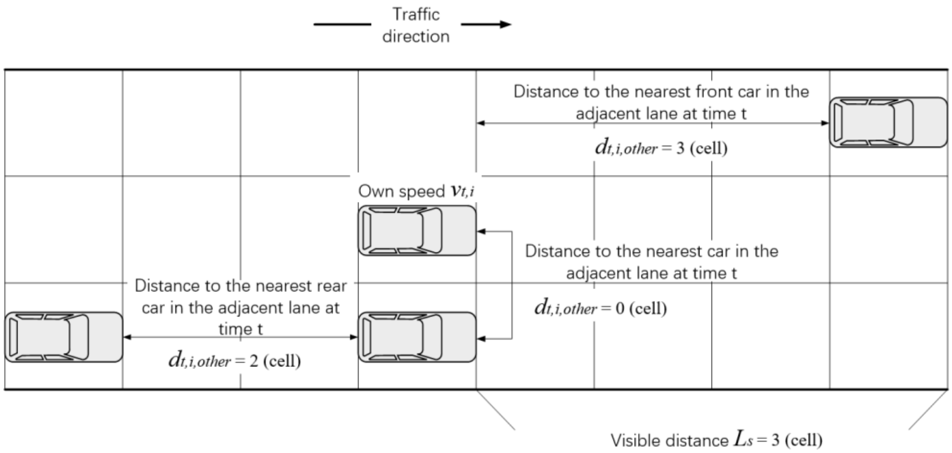


Figure 3. Simulation diagram of the distance between the vehicle and adjacent vehicles in the state of dense fog.

(1) Determine the random braking probability P_{brake}

$$\Delta v_{t,i,other} = \min(v_{t,i} - v_{t,i+1,other}) \tag{11}$$

$$P_{brake} = \begin{cases} \max\left\{0, P_0 \cdot \left[1 + \gamma \tanh\left(\left|L_s - d_{t,i,other}\right|\right) + \delta \frac{\Delta v_{t,i,other}}{v_{\max}}\right]\right\}, 0 \le d_{t,i,other} \le L_s \\ P_0, d_{t,i,other} > L_s \end{cases} \tag{12}$$

where the random braking probability P_{brake} is influenced by the vehicles in the adjacent lane; P_0 is the random braking probability of the vehicles in a dense fog environment that are not influenced by the adjacent lane; the speed of the vehicle i at time t is denoted as $v_{t,i}$, the speed of the nearest preceding vehicle i in the adjacent lane at time t is denoted as $v_{t,i+1,other}$; the maximum speed of the vehicle is $v_{\max}$; $\Delta v_{t,i,other}$ is the speed difference between the first vehicle and the nearest preceding vehicle in the adjacent lane at time t (cell/s); and $d_{t,i,other}$ is the distance between the first vehicle and the nearest preceding vehicle in the adjacent lane at the time (cell); $\gamma = 0.4$, $\delta = 0.5$.

(2) Vehicle state

1. Acceleration

$$v_{t,i} = \min(v_{t,i} + 1, \alpha_n \cdot v_{\max}) \tag{13}$$

2. Random braking The driver has the probability of braking in a small area during the nonacceleration and deceleration states.

$$v_{t,i} = \begin{cases} \max(v_{t,i}, 0), d_{t,i} > L_s \\ \max(v_{t,i} - 1, 0), 0 \le d_{t,i} \le L_s \end{cases} \quad with P_{brake} \tag{14}$$

3. Forced braking

$$v_{t,i} = \min(v_{t,i}, d_{t,i} - 1) \tag{15}$$

4. Position update

$$x_{t+1,i} = x_{t,i} + v_{t,i} \times \Delta t \tag{16}$$

2.4. Determination of Model Key Parameters

According to the SIR model in Section 2.1, vehicles in congestion affect the remaining uncongested vehicles in the surrounding area by random probability with λ as the average propagation rate. At the same time, the vehicles move away from the congested area with the average recovery rate of μ ($I \to R$ transition). The vehicles affected by the congested vehicles also become congested vehicles through the average propagation rate λ ($S \to I$ transition). The state of the vehicles themselves changes as shown in Figure 4.

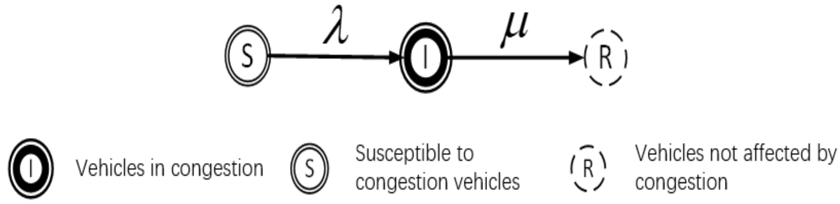


Figure 4. Vehicle changes according to SIR congestion propagation model.

According to the CA-SIR model in Section 2.2, the congestion propagation probability λ_{CA} and congestion recovery probability μ_{CA} are respectively related to different braking probabilities P_{brake} and lane-change probabilities P_{change}. By quantifying them and adding the speed limit requirements for different visibilities in a foggy environment, the congestion propagation probability λ_{CA} and congestion recovery probability μ_{CA} in this scenario are obtained as follows:

$$\lambda_{CA} = \frac{(1-\alpha_i) \times \overline{\lambda} \times \delta}{20 \times e^{-P_{change} \times P_{brake}}} \quad (17)$$

$$\mu_{CA} = \frac{(1-\alpha_i) \times \overline{\mu}}{20 \times e^{-P_{change} \times P_{brake}}} \quad (18)$$

where $\overline{\lambda}$ indicates the average congestion propagation probability, $\overline{\lambda} = 0.75$; $\overline{\mu}$ is the average congestion recovery probability, $\overline{\mu} = 0.4$; and the speed factor α_i in the foggy environment is affected according to different visibilities and takes a value between 0 and 1, and the equation is

$$\alpha_i = \frac{v_{i,\max}}{v_{\max}} \quad (19)$$

where $v_{\max}$ = 7 cell/s is the maximum speed in clear weather. $v_{i,\max}$ is the speed limit value in different foggy weather environments.

3. Results and Discussion

3.1. Assumptions and Parameters

According to Tables 1 and 2, the visibility environment in this study was divided into four types: light fog, medium fog, heavy fog, and dense fog, with different visibility distances under different visibilities and different degrees of influence on the driver's vision. The model's speed limit rules under each type were adopted from Table 1 to obtain the maximum speed of the vehicle (light fog), (light fog), (heavy fog), and (dense fog). The values of speed coefficients for different visibilities were 1, 0.75, 0.5, and 0.25, respectively.

Through the MATLAB simulation study, it was found that the model simulation is not feasible when the braking probability is too large under all four fog types. The vehicle causes or encounters congested conditions once it is driven out, the effect of the random Poisson distribution generation of the vehicle is invalid, and the model complies with the maximum speed limit requirement in low visibility with low overall vehicle speed. Therefore, once the driver brakes randomly on the vehicle in the slow speed state, the vehicle speed goes to zero while the impact range is larger. Vehicles affected by congestion due to foggy weather congestion, after starting to accelerate, also need to maintain the minimum safety distance, while the recovery speed is slow and can easily cause comprehensive traffic paralysis. Therefore, according to the relevant research [13], the set braking probabilities were 0.24 (light fog), 0.26 (medium fog), and 0.28 (heavy fog). The random braking probability for vehicles not affected by adjacent lanes in dense fog was 0.31 [20]. The lower the visibility, the lower the driver's willingness to change lanes, and the probabilities of changing lanes were set to 0.41 (light fog), 0.23 (medium fog), and 0.14 (heavy fog) for each of the three cases. In the case of dense fog, the driver's individual lane change was ignored.

3.2. Road Time–Space Change Graph under Foggy Environment

Due to the different minimum safety distances under different visibilities, in order to avoid the vehicles at the beginning of each lane maintaining the distance and causing congestion, and at the same time, considering that vehicles driving under foggy conditions are affected by the speed limit, the total time of the overall simulation is too long. Therefore, only the spatial and temporal changes of the road in the middle period were intercepted for research and analysis.

Figure 5 shows the time–space change graph of some vehicle with a visibility of 400 m under a light fog condition, where the vehicle traveled 1721 s in 500 cells, and 600–1200 s was selected as the main reference time in the model simulation process. In Figure 5, the straight line formed by the points of the same vehicle at different position–time points represents the speed of the vehicle (black arrow line), and the slope of the line represents the rate of the vehicle. The larger the slope, the slower the vehicle travels, and the arrow marks the direction the vehicle travels. From Figure 5, it can be found that the first vehicle to exit (the vehicle below the black arrow line) is less likely to be affected by the vehicle in front of it and travels at the highest rate at that visibility. The denser the black area, the higher the density of vehicles, which can quickly dissipate by changing lanes or accelerating in that visibility, avoiding traffic congestion.

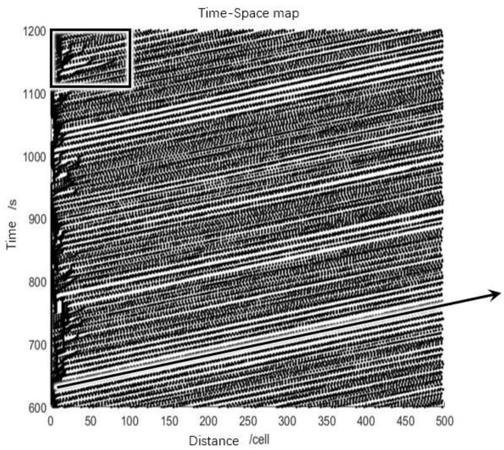


Figure 5. Vehicle time–space variation at 400 m visibility (light fog).

The denser traffic flow at the starting point is due to forced braking by vehicles needing to maintain a minimum safe distance from the vehicle in front, resulting in a brief gathering of vehicles at a nearby location. The maximum queue caused by the congestion is shown in Figure 5 in the black box, with a length of about 100 m and a duration of about 30 s, which dissipated on its own. After the distance and speed adjustment of the vehicles at the starting point, the subsequent simulation process was a not-obvious crowding process.

Figure 6 shows the time–space variation of some vehicles at 170 m visibility in medium fog, and the same 400–800 s range was selected as the main reference duration during the model simulation. The vertical line in the black box in the figure, which is different from the vehicle trajectory, represents the congestion phenomenon at this location. The gray arrow line represents the average speed of the vehicle at that visibility without the effect of congestion. The model simulations for the same number of vehicles lasted for a total of 1296 s. Comparing the information in Figure 6 with that in Figure 5, it can be observed that:

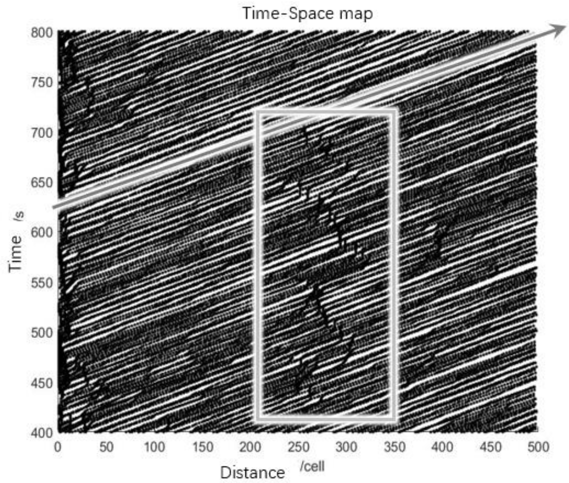


Figure 6. Vehicle time–space variation at visibility of 170 m (medium fog).

(1) The slope of the gray arrow line in Figure 6 is steeper than the slope of the black arrow line in Figure 5, which indicates that the average speed of vehicles under medium fog without the influence of congestion is lower than that under light fog, which was 16.03 m/s (light fog) and 8.63 m/s (medium fog), respectively, and the reduced visibility brought about reduced sight distance and reduced speed limit of the highway. The difference between the speed limit ratio of 25% for both vehicles and the average speed ratio of 46.16% for both vehicles was obvious, which indicates that the congestion phenomenon under the medium fog condition is characterized by a small range but a high frequency.

(2) Congestion in the medium fog state occurred frequently, but almost all appeared in the gray box (the 200th metric cell to the 300th metric cell); further, there was no obvious congestion propagation phenomenon, the overall frequency of congestion occurrence was related to the random braking probability in this visibility, and overall, there was no obvious pattern.

(3) The total elapsed time for the same number of vehicles traveling the same distance at 170 m visibility (1296 s) was inversely reduced compared with the total simulation time at 400 m visibility (1721 s), and the difference accounted for 24.7%, which was not caused by errors. By constantly changing the visibility in the light fog range, the overall simulation time was around 1680–1780 s, which was much higher than that in the medium fog condition. By analyzing the simulation content, it was found that the speed limit in accordance with the Road Traffic Safety Law was too low in the light fog state, and at the same time, vehicles did not hesitate to slow down in order to maintain the minimum safety distance, thus causing congestion, and although the congestion range was not large and could dissipate by itself, the overall time spent was longer.

Figure 7 shows the time–space changes of some vehicles in the heavy foggy condition when the visibility is 75 m. In order to clearly see the spatial and temporal diagram of the vehicles driving in this visibility, we intercepted and enlarged the model simulation process of 500–700 s as the main reference time. The black arrow line represents the speed of the vehicles in this figure when they are not affected by congestion, and the model simulation with the same number of vehicles at a safe distance lasted a total of 1234 s. Comparing Figure 7 with Figures 5 and 6, it can be found that:

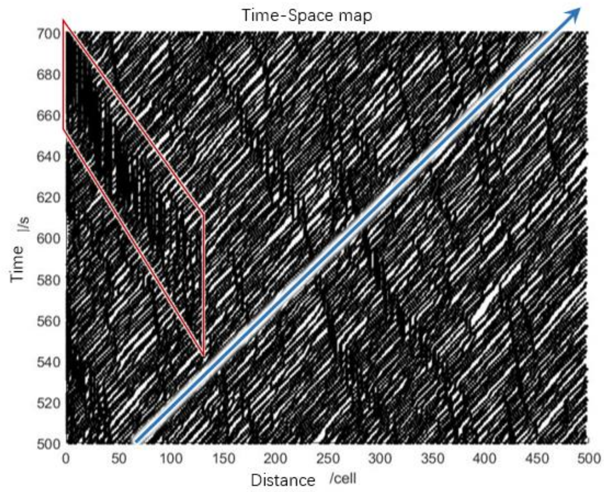


Figure 7. Vehicle time–space variation at visibility of 75 m (heavy fog).

(1) The slope of the blue arrow line in Figure 7 is greater, and the average vehicle speed when the visibility is 75 m was 3.28 m/s when not affected by congestion, which is a 62% decrease compared to a medium fog day. It is mainly because the foggy condition is more severely affected by visibility and the specified speed limit is lower.

(2) The congestion propagation phenomenon in the heavy fog condition was obvious, and as shown in the red area of Figure 7, the congestion range gradually moved upstream with the passage of time and became denser and denser. In addition, due to the reduced visibility and the influence of the speed limit, different driving styles of drivers brought obvious differences, with conservative drivers driving slower, braking more easily, and staying longer in the lane. It can cause traffic disorder and congestion to subsequent vehicles. At the same time, congestion propagation did not dissipate on its own in low visibility.

Therefore, in the heavy foggy condition, the congestion continues all the time, which means that it remains in the I state and does not reach the R state, much less return from the R state to the S or I state. The propagation of congestion is just a movement in space and does not involve passing or feedback between states. Therefore, to ease the traffic congestion phenomenon is the primary goal in a heavy fog environment on the freeway.

(3) The total elapsed time for the same number of vehicles traveling the same distance at 75 m visibility (1234 s) was about the same as the total simulation time at 170 m visibility on a medium fog day (1296), with a difference of less than 5%. This indicates that although the number of vehicles crowded in heavy fog is significantly greater than in medium fog, the average speed is also slower. However, the number of vehicles in heavy fog that needs to maintain the minimum safety distance was also smaller; therefore, despite the slower speed, the vehicle density was greater, and as a result, the overall time used in the simulation was almost the same.

In the vehicle time–space change map in Figure 8 for the dense fog state with a visibility of 40 m, the vehicle speed under the fog environment was very slow due to the speed limit; therefore, the entire model simulation time for the same number of vehicles lasted a total of 1962 s, and by intercepting and enlarging the model simulation process using 800–1100 s as the main reference time, we can see more clearly the time–space map of the vehicle driving under a given visibility. The red arrow line represents the speed of the vehicle when it is not affected by congestion.

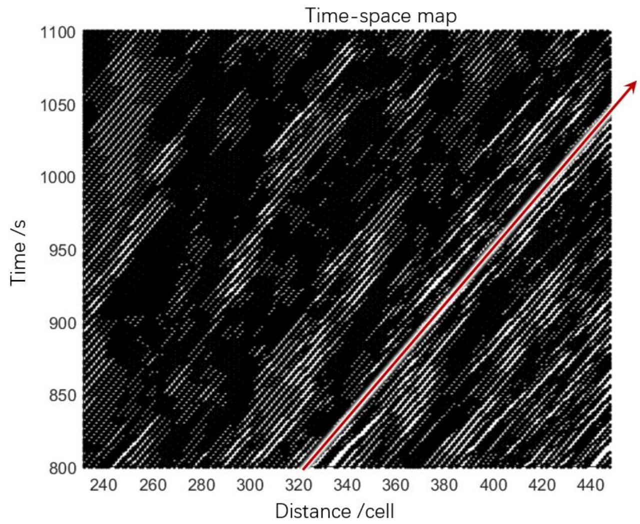


Figure 8. Vehicle time–space change map at visibility of 40 m (dense fog).

Comparing Figure 8 with Figures 5–7, according to the vehicle time–space variation diagrams under the four visibility levels, it can be found that:

(1) Due to the different maximum driving speeds, visible distances, and minimum safety distances in each visibility, the slope of the arrow line in Figure 8 gradually becomes larger, representing that the vehicle speed in the model gradually slows down even if it is not affected by congestion. On a dense fog day, the average speed was 2.51 m/s.

The average speed obtained from the simulation in all four foggy conditions without congestion was well below the maximum speed limit specified for the visibility, averaging only 55% of it. The most obvious was that in the heavy fog environment, at only 32.8% of the maximum speed limit. It can be seen that a foggy environment in general has an impact on the speed of vehicles, in addition to the dense fog days, and the reduction in visibility on the impact of speed gradually increased. On dense fog days, the general driving behavior of vehicles was more conservative, and the nearest vehicles within the visibility range tended to take the way of companionship, which could effectively reduce the risk generated by congestion under the speed limit conditions.

(2) The illustrations of the congestion propagation phenomenon in Figures 7 and 8 show that the biggest difference is that the congestion propagation in the foggy condition gradually moved upstream with time and the congestion range increased, while in the dense fog environment, when the traffic flow was too dense or the previous vehicle suddenly braked and caused congestion, the congestion range around the vehicle moved downstream synchronously with the passage of time. In other words, under heavy fog, vehicles in congestion move away, and new vehicles move into the congestion area upstream; however, under dense fog, vehicles in congestion move forward together with the surrounding vehicles at a slow speed and do not move away from the congestion area.

Therefore, in that visibility, in addition to speed limit measures, requiring vehicles in the nearest ramp to leave the highway as soon as possible, strictly controlling the density of vehicles, or implementing the intermittent release of multiple lanes is necessary to avoid causing congestion pile-ups or the traffic paralysis phenomenon, which is more likely to cause traffic accidents in low visibility.

3.3. Analysis of Speed Characteristics in Foggy Environment

Figure 9 shows the overall average vehicle speed for each visibility condition, and the vehicle speed in the model is in cell/s.

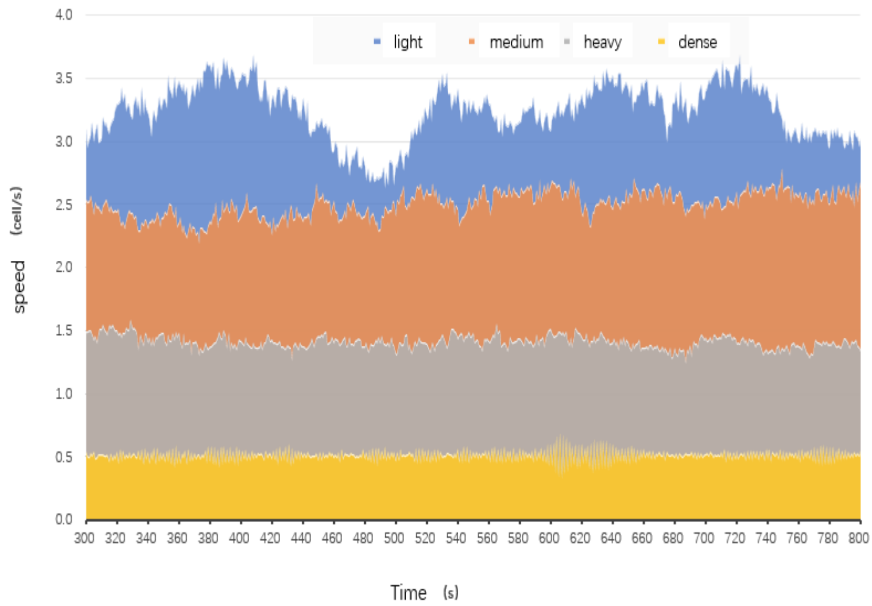


Figure 9. Overall average speed of vehicles at various visibility levels.

From this, it can be seen that:

The average vehicle speed fluctuated the most on light fog days, and as the visibility decreased, the average travel speed also decreased, and the up and down fluctuation decreased.

In the light fog environment, the overall speed variation was as low as 2.66 cell/s and as high as 3.7 cell/s. Combined with the time–space map in Figure 5 when the visibility is 400 m, it can be found that the fluctuation of the overall average vehicle speed is consistent with the sparse and dense distribution of traffic flow in Figure 5, which indicates that in a high-visibility environment, vehicles driving on the highway are still mainly influenced by the traffic flow, similar to in sunny weather, and although there is frequent acceleration and braking behavior, it is not too affected by the weather and almost not affected by the sight distance factor brought by light fog.

As visibility gradually decreased, the overall vehicle speed decreased and fluctuated to a smaller extent. The overall average vehicle speed began to be affected by the reduced visibility and speed limits. The overall traffic flow density gradually became greater, especially in dense fog conditions, and the travel speed was the slowest, causing an increase in traffic flow and a greater range of congestion and queuing vehicles. In addition, the visibility was reduced, the driver's field of vision was also reduced, and the greater impact of the vehicle in front and the surrounding vehicles was more likely to cause congestion propagation.

Therefore, in a foggy environment, freeway authorities should formulate the best speed limit value under different visibilities. If necessary, freeway authorities need to strictly control the amount of vehicle on the freeway. Even for high-density traffic in low visibility, highway authorities need to implement the intermittent release of different lanes of traffic to avoid the spread of congestion. In this case, shared mobility, such as bicycle sharing, can be considered to decrease vehicle volume and reduce congestion to improve safety [27,28].

4. Conclusions

For the foggy scenarios of freeways under different visibilities, this study constructed a CA-SIR cellular automata epidemic model to study this congestion propagation regulation. In the model construction, the key parameters in CA-SIR were determined by establishing the car-following and lane-changing rules for two foggy scenarios. Finally, the evolution of congested vehicles at each time point of congestion propagation on the freeway under a foggy environment was obtained.

Through related simulation analysis, it was found that as the visibility decreases and the vehicle speed decreases, the congestion propagation phenomenon gradually becomes prominent, and the congestion queuing phenomenon becomes more obvious. In light fog conditions, the speed limit is an important factor causing traffic congestion; in medium fog conditions, the amount of congestion generated is related to the random braking probability set by the vehicles; and in heavy fog conditions, the congestion area gradually moves upstream while vehicles keep driving away from the congestion area. In dense fog, affected by very low visibility, vehicles are more likely to choose to travel in groups, and the congestion range keeps moving downstream in parallel with the passage of time. It is difficult for vehicles to leave once they are affected by congestion, and the congestion does not dissipate easily.

This study mainly investigated freeway congestion propagation in foggy weather from the perspective of traffic flow. Moreover, the reduced visibility in an actual foggy weather environment may significantly affect drivers' psychology and related visual exhaustion and driving risk perception. These factors need to be further explored in depth and can be subsequently considered to be integrated into the model of congestion propagation.

Author Contributions: Conceptualization, J.Y.; methodology, J.H.; validation, Y.B.; formal analysis, J.H.; investigation, J.L. and Y.H.; resources, J.Y.; data curation, J.H.; writing—original draft preparation, J.H. and Y.B.; writing—review and editing, J.Y.; visualization, J.H.; supervision, Y.H.; project administration, Y.H.; funding acquisition, J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by Ministry of Education in China (MOE) Project of Humanities and Social Sciences (22YJAZH131) "Research on mechanism of route coordinated control of commuter traffic at the urban road network in the environment of big data". The authors, therefore, acknowledge with thanks the financial support of Ministry of Education in China (MOE).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, X.C.; Chen, Z. *Data-Driven Freeway Performance Evaluation Framework for Project Prioritization and Decision Making*; MPC-17-316; Mountain-Plains Consortium, North Dakota State University, Upper Great Plains Transportation Institute: Fargo, ND, USA, 2017.
2. Chris, W.; Penina, R. The conceptual structure of traffic jams. *Transp. Policy* **1998**, *5*, 23–35.
3. Orosz, G.; Wilson, R.E.; Szalai, R.; Stépán, G. Exciting traffic jams: Nonlinear phenomena behind traffic jam formation on highways. *Phys. Rev. E Statal Nonlinear Soft Matter Phys.* **2009**, *80*, 046205. [CrossRef] [PubMed]
4. Schrank, D.; Albert, L.; Eisele, B.; Lomax, T. *2021 Urban Mobility Report*; Texas A&M Transportation Institute: Bryan, TX, USA, 2021.
5. Ministry of Transport Public Bureau. *2019 China Highway Network Operation Blue Book*; People's Transport Publishing House: Beijing, China, 2020.
6. Druta, C.; Kassing, A. Assessing driver behavior using shrp2 adverse weather data. *J. Saf. Research.* **2020**, *73*, 283–295. [CrossRef] [PubMed]
7. Zheng, S.; Li, Z.; Zhao, X. Behavioral Characteristics of Lane-changing on Highway in Foggy Weather. *J. Transp. Inf. Saf.* **2020**, *38*, 35–42, 51. [CrossRef]
8. Yan, X.; Li, X.; Liu, Y.; Zhao, J. Effects of foggy conditions on drivers' speed control behaviors at different risk levels. *Saf. Sci.* **2014**, *68*, 275–287. [CrossRef]
9. Zhang, X.; Gao, J.; Liao, L. Car Following Behavior of an Expressway Driver in Fog Environment. *China J. Highw. Transp.* **2022**, *35*, 275–285. [CrossRef]
10. FHWA Road Weather Management. How Do Weather Events Impact Roads? 2018. Available online: https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm (accessed on 21 July 2022).
11. Brooks, J.O.; Crisler, M.C.; Klein, N.; Goodenough, R.; Beeco, R.W.; Guirl, C.; Tyler, P.J.; Hilpert, A.; Miller, Y.; Grygier, J.; et al. Speed choice and driving performance in simulated foggy conditions. *Accid. Anal. Prev.* **2011**, *43*, 698–705. [CrossRef]
12. Hassan, H.M.; Abdel-Aty, M.A.; Choi, K.; Algadhi, S.A. Driver Behavior and Preferences for Changeable Message Signs and Variable Speed Limits in Reduced Visibility Conditions. *J. Intell. Transp. Syst.* **2012**, *16*, 132–146. [CrossRef]

13. Ahmed, M.; Ghasemzadeh, A.; Eldeeb, H.; Gaweesh, S.; Clapp, J.; Ksaibati, K.; Young, R.K. *Driver Performance and Behavior in Adverse Weather Conditions: An Investigation Using the SHRP2 Naturalistic Driving Study Data—Phase 1*; Department of Civil and Architectural Engineering, University of Wyoming: Laramie, WY, USA, 2015.
14. Liu, Z.; Yang, X.; Wu, X.; Huang, Z. Modeling and Simulation of Car Following in Fog Based on Cellular Automata. *J. Syst. Simul.* **2021**, *33*, 2399–2410. [CrossRef]
15. Liu, Z.; Yu, C.; Wang, C.; Li, Q. Influence of Foggy Environment on Expressway Car-following Safety. *J. Chongqing Jiaotong Univ. Nat. Sci.* **2019**, *38*, 88–94. [CrossRef]
16. Feng, W.; Zhang, C.; Cao, Y.; Zhang, T.; Luo, S. Modeling and Analysis of Traffic State of Regional Freeway Network in Fog Region. *J. Wuhan Univ. Technol. Transp. Sci. Eng.* **2021**, *45*, 93–98. [CrossRef]
17. Shah, S.; Mansoor, M.; Mirza, A.; Dilshad, M.; Khan, M.U.; Farwa, R.; Khan, M.A.; Bilal, M.; Iqbal, H.M.N. Predicting COVID 19 Spread in Pakistan using the SIR Model. *J. Pure Appl. Microbiol.* **2021**, *15*, 462–463. [CrossRef]
18. Miguel, L.; Alberto, P.; Andrés, O. An extensive validation of a SIR epidemic model to study the propagation of jamming attacks against IoT wireless networks. *Comput. Netw.* **2019**, *165*, 106945. [CrossRef]
19. Saeedmanesh, M.; Geroliminis, N. Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. *Transp. Res. Part B Methodol.* **2017**, *105*, 193–211. [CrossRef]
20. Hofleitner, A.; Herring, R.; Bayen, A.; Han, Y.; Moutarde, F.; de La Fortelle, A. Large scale estimation of arterial traffic and structural analysis of traffic patterns using probe vehicles. *Transp. Res. Board Meet.* **2012**, 1–22. [CrossRef]
21. Zheng, S.; Han, X. A Simulation of Cellular Automata Based on the SIR Infectious Disease Model with Multifactorial Constraints. *J. Guangdong Univ. Technol.* **2018**, *35*, 51–59. [CrossRef]
22. Pellis, L.; Ferguson, N.M.; Fraser, C. The relationship between real-time and discrete-generation models of epidemic spread. *Math. Biosci.* **2008**, *216*, 63–70. [CrossRef]
23. Ji, H.; Wang, Y.; Li, P.; Su, B. Traffic flow simulation of urban three-lane road considering influence of accident vehicle. *China Saf. Sci. J.* **2021**, *31*, 112–120. [CrossRef]
24. Xiao, J.; Qin, Y.; Wang, Y.; Li, H. Research on the Relationship between the Driver's Temperament and Driving Behavior. *Chin. J. Ergon.* **2014**, *20*, 23–27. [CrossRef]
25. Yu, L.; Zhang, M.; Wang, X. Simulation Study on Safety Speed Limit of Fog Weather Based on Car Following Model. *Sci. Technol. Eng.* **2018**, *18*, 224–229. [CrossRef]
26. Tan, J. *Rear-End Collision Risk Management of Freeways in Heavy Fog*; Tsinghua University: Beijing, China, 2015.
27. Turoń, K.; Czech, P.; Tóth, J. Safety and security aspects in shared mobility systems. *Sci. J. Sil. Univ. Technol. Ser. Transp.* **2019**, *104*, 169–175. [CrossRef]
28. Turoń, K.; Sierpiński, G. Bike-sharing as a possibility to support Vision Zero. *MATEC Web Conf.* **2018**, *231*, 03005. [CrossRef]

sustainability

Article

Transit-Oriented Development in Saudi Arabia: Riyadh as a Case Study

Khalid Mohammed Almatar

Department of Urban and Regional Planning, College of Architecture and Planning, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31451, Saudi Arabia; kmalmatar@iau.edu.sa

Abstract: Transit-oriented development (TOD) in Saudi Arabia is becoming a significant priority for the government and developers to create a sustainable and quality living environment. TOD is an integrated transport and urban planning method that aims to reduce car use and urban sprawl, increase the use of public transport, and enhance sustainable mobility. To meet the global goals as per the Paris Accord, Saudi Arabia's policymakers must prioritize the integration of TOD in urban planning. This study was carried out with the main aim of identifying the environmental, social, and economic benefits of implementing TOD in Riyadh, Saudi Arabia. A mixed-study research method was used, and data were collected using a questionnaire survey and semi-structured interviews. The quantitative data were analyzed using SPSS version 21, and qualitative data were analyzed using NVivo software. The findings of this study show that TOD in Riyadh City would positively impact economic, environmental, and social aspects. TOD would reduce travel time, allow its people to have an active lifestyle, and reduce congestion. TOD would help reduce mental health disorders and improve physical activity. TOD would positively impact the environment of Riyadh City and assist in reducing greenhouse gases. Overall, the study results provide a reliable perspective on the benefits of TOD. Most participants assumed that the implementation of TOD in Riyadh City would increase automobile mobility, provide more employment opportunities, and reduce travel time, positively impacting the environment and economy of Riyadh City.

Keywords: transit-oriented development; sustainable transportation; TOD; Riyadh; KSA

check for updates

Citation: Almatar, K.M. Transit-Oriented Development in Saudi Arabia: Riyadh as a Case Study. *Sustainability* **2022**, *14*, 16129. https://doi.org/10.3390/su142316129

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 24 September 2022
Accepted: 28 November 2022
Published: 2 December 2022

1. Background

Over the last few decades, the significant increase in private vehicles has led to chronic traffic congestion and has become a substantial problem due to the impact on economic growth, air quality, and sustainable living [1]. The transportation system is the primary source of urbanization and infrastructure development and is essential for a country's economic growth and viability [2]. The development of a public transportation system is vital for a sustainable environment. Transport is one of the highest energy-demanding sectors and a leading contributor to greenhouse gas emissions. In 2019, the GCC countries contributed to approximately 18% of total carbon emissions [3]. In Saudi Arabia, the transport sector contributes 22% of the overall carbon emissions [4]. The development of a robust transportation system and its associated policy is beneficial for the entire environment, and most countries across the globe are focusing on it to bring more attention to it. However, due to the high growth of these economies, public health has now emerged as a significant concern for these countries. Hence, eco-friendly transportation has become an integral and significant part of the development of communities and is emerging as a significant element in overall economic well-being [5]. Recently, transit-oriented development (TOD) has been given great importance and is considered by many countries worldwide for achieving environmental sustainability [6]. TOD is generally defined as an integrated method of land and transport use planning that makes transit use, cycling, and walking desirable and convenient and maximizes the existing transport system efficiency by concentrating on development. TOD aims to produce pedestrian-friendly and practical urban neighborhoods, incorporating schools, lodging, travel, and other monetary and social improvements, thus benefiting different linked groups accordingly [6].

Saudi Arabia is one of the most urbanized countries in the world, as 80% of the country's population lives in urban areas [7]. Due to high urbanization and development, transportation usage has been significantly increasing, increasing environmental problems and the consumption level. In 2017, the energy consumption from the Saudi Arabian transport sector was almost 20%, mainly based on fossil fuels [8]. TOD gained immense appreciation, especially in Riyadh, Saudi Arabia, because of the recent development in the city. With the expectation to reduce carbon emissions, the Riyadh metro system was developed to combat climate change and transportation-related carbon emissions. In addition, Riyadh is also implementing a bus network and changing the urban framework through transit-oriented development [9]. The most significant advantage of this integrated public system, or TOD, is reducing the per capita energy consumption in the city and ultimately reduce traffic congestion, which has a direct and positive impact on the economic and environmental sustainability of the country. Transit development aims to attain the core production of practical, effective, and blended neighborhoods that incorporate a high and efficient level of travelling [9]. Lodging, social parks, and other significant monetary enhancements are some of the significant elements covered under the umbrella of TOD [10]. One of the most important things associated with TOD is its arrangements in urban sprawl and arranging systems for green building scenarios by providing access to make transportation more viable and competitive [10]. In other words, TOD is an initiative for Riyadh City to move toward greener options and become competitive and feasible. This study was carried out with the main aim of determining the benefit and need of transit-oriented development in Saudi Arabia, with a specific focus on Riyadh. This study will help the government of Riyadh make informed decisions about the development and planning of the transport system.

2. Literature Review

Peter Calthorpe introduced the TOD concept with a specific focus on transit planning [11]. TOD has mainly been categorized as one kind of sustainable urban transport that creates a high level of human interaction. Bernick and Cervero have strongly focused on the three Ds' (design, diversity, and density) role in TOD [11,12]. During the 1980s, after observing suburban congestion, researchers and urban designers focused on alternatives, such as developments around transit areas, with the main objective of reducing motorized trips. During the Paris Agreement in 2015, the kingdom (KSA) showed its commitment to reducing the emissions of greenhouse gases and adhering to the mitigations related to climate change [13]. The kingdom has demonstrated its dedication and capacity to reduce more than 130 million tons of carbon dioxide by the end of 2030 [13]. The vision of 2030 is set out in the country's national determined contribution (NDC), which is aligned with the United Nations' concept on the issue of climate change. This NDC can bring a high level of economic diversification and become highly adaptive to the impact of climate change [13,14]. According to Noland et al. (2017), both transportation and infrastructure are considered significant forms of urbanization [15]. In the current economic times, urban-based transportation is gaining close attention. It is a significant pillar for travelers' mobility that focuses specifically on expanding these areas. Statistics revealed that 80% of the country's population lives in urban areas, which shows that urbanization is at its peak in the country [15]. Papa and Bertolini (2015) indicated that since Riyadh is becoming a financial hub for Saudi Arabia, and the current king is looking forward to expanding its position in the context of the goal for 2030, the congestion problem may increase rapidly [16]. Thus, the country desperately needs TOD to achieve sustainable transportation that aims to minimize or eradicate environmental problems, maximizing the quality of life by enhancing social-based inclusions. Wey and Chiu (2013) conducted a study in which the benefits related to TOD were mentioned [17]. The study was conducted in Turkey, in which the sustainable tourism factor was highlighted. The researcher used quantitative measures to complete the research, using quantification for the same purpose. The researcher used a sample of 404 individuals and found that countries can achieve social benefits by implementing TOD. The analysis found that social indicators include the status of human health, community-based livability, and the equity factor. The researcher revealed that TOD is helpful for the people of Turkey to obtain affordable, green, and sustainable transport that could positively increase their social well-being [17].

There was another study conducted in the same domain, and the research was conducted by Renne (2009). The researcher identified that one of the main benefits of TOD is the achievement of environmental sustainability using environmental indicators. Environmental sustainability includes overcoming pollution, creating a better environment for wildlife, and creating a manageable vehicle system. These elements are achievable, and the researcher has identified them before [18]. Endorsing the same idea, Zandiatashbar, Hamidi, Foster, and Park (2019) revealed that, apart from environmental indicators, certain indicators of economic well-being are directly and indirectly connected with TOD [19]. These economic indicators help gain immense benefits and viability in the transportation system and benefit the overall community [19]. The studies covered above were generally conducted in the European areas, and very few studies have been conducted specifically in Saudi Arabia, especially in Riyadh. Therefore, this study was carried out to determine the benefit of TOD in Riyadh City.

3. Methods

3.1. Study Setting

In the Arabian Peninsula, the Saudi Arabian kingdom is the largest nation, with a population of more than 30 million and an area of 2.14 million km^2. The geographical focus of this research study was Riyadh City, the capital of Saudi Arabia, with a population of 5.5 million (Figure 1). Currently, the spatial area of Riyadh City is almost 3000 km, covering 209 districts and 13 municipalities. In 2014, the city initiated the integrated transport system, with a budget of more than USD 23 billion. This was important, as the city's population is expected to reach 10 million by 2030.

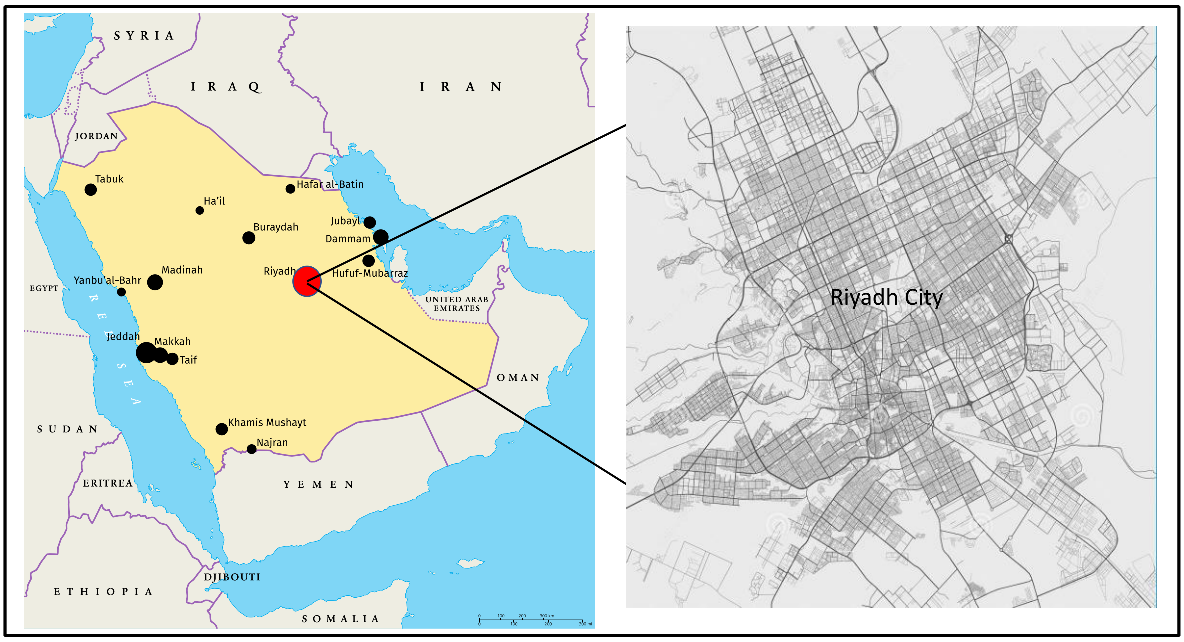


Figure 1. Riyadh City Map.

3.2. Research Method

The focus of this study was to evaluate the benefits of transit-oriented development in Riyadh, Saudi Arabia. For the present study, both qualitative and quantitative methods were adopted to determine the TOD benefit in Riyadh City.

3.3. Data Collection

The data were collected in two phases. In phase 1, quantitative data were collected using the questionnaire. In phase 2, qualitative data were collected by conducting interviews. Purposeful sampling was used based on the participant's willingness to participate in the questionnaire survey. The questionnaire was distributed among 120 individuals, among which only 95 were duly filled. Five in-person semi-structured interviews were conducted for the qualitative data to obtain valid and reliable data related to their research objectives, as shown in Figure 2. In addition, this allowed researchers to collect more complex data and comprehensively represent the research area. Semi-structured interviews were conducted with stakeholders, including transport experts, financial consultants, and environmental specialists. The semi-structured interview comprised unstructured and structured interviews where themes and question lists were used to collect data.

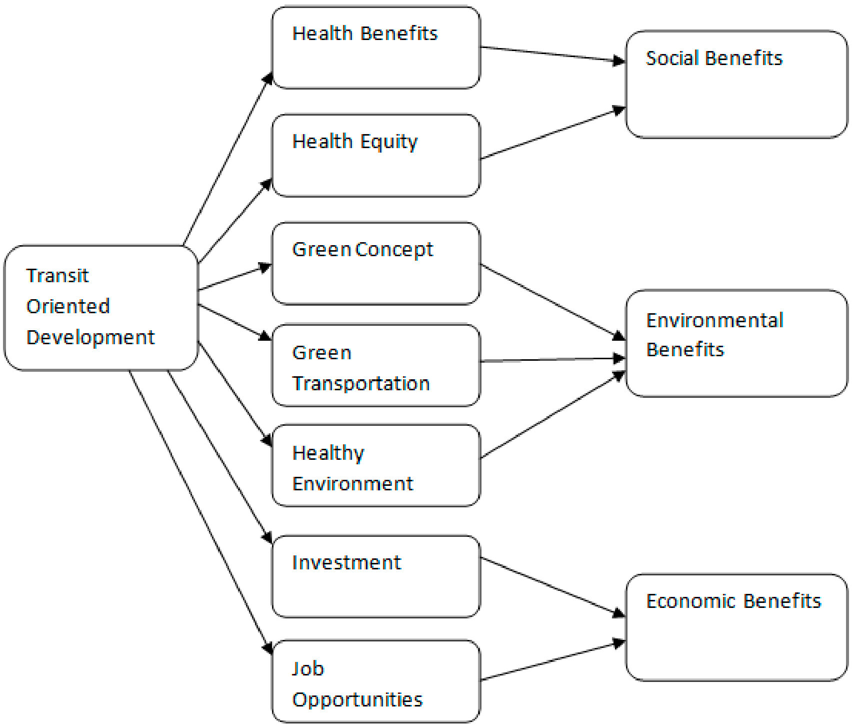


Figure 2. Framework for the Data Collection.

Protocol of Interview

The interviews were carried out in person. Before conducting the interviews, the documents related to the interviews were sent to all participants via hardcopy or email. All participants were asked the same questions; each interview was 30–40 min long. After providing the participants with the details of the study, permission was obtained to record the interview. In this interview, closed and open-ended questions were used to address the research question and to have appropriate data addressing the research objectives. The interview mainly focused on the selected aspect of TOD, as shown in Table 1. Thus, participants were requested to demonstrate to what extent these aspects would be beneficial by implementing TOD intervention in Riyadh City.

Table 1. Protocol of Interview.

Section	Purpose	Contents
Opening	Details of the research study objectives are provided to the participants, and permission is taken to record the interview.	Start of the interview/Rapport is built
1	To investigate the benefits of TOD in terms of economic, environmental, and social benefits.	What benefits do you think TOD implementation will provide in terms of economic, environmental, and social aspects?
2	Measure participant's opinions for the necessity of implementing TOD in Riyadh City.	To what extent do you agree or disagree with the implementation of TOD in Riyadh City?

3.4. Data Analysis

The qualitative data collected from participants were transcribed and then documented. The NVivo software was used, and analysis was performed using grounded theory. Based on the study question designs and results interpretation, the 'majority' or 'most' of the participants here indicate 75% or more participants with a similar belief. 'Many' refers to 75 and 50%, 'some' indicates between 25 and 50% of the participants having similar beliefs, and 'few' indicates less than 25% of the participants. The quantitative data were analyzed by using SPSS version 21 to determine the mean score.

4. Results

The main aim of this section is to present the findings to address the research question. This section presents information on the benefits of transit-oriented development in Riyadh, Saudi Arabia.

4.1. Semi-Structured Interview Findings

4.1.1. Economic Benefits

Most participants stated that transit-oriented development would add significant economic value to the economy of Riyadh City. One participant indicated that TOD is expected to reduce unemployment by providing job opportunities for Saudi citizens and would improve city mobility. Most participants stated that presently, the roads of Riyadh City are over-congested, which negatively impacts household mobility and businesses, so they hoped TOD would reduce urban sprawl and increase public transport efficiency. Another participant stated that transit ridership would increase by implementing TOD, further increasing fare revenues and bringing additional revenues into the transit system and localities. Most participants stated that TOD would increase public transportation usage, leading to the economic efficiencies of these systems and improving the city's business productivity. Few participants stated that TOD would increase property value. Most participants indicated that TOD would help reduce transportation and infrastructure costs, such as garbage pickup, the cleaning of roads, and the reduction in external fuel cost consumption and air pollution.

4.1.2. Environmental Benefits

In Saudi Arabia, a high flow of motor vehicles has been observed due to economic growth and increased urbanization. Most of the participants in this study stated that transit development would help attain benefits to the quality of the urban environment, such as increased public transportation usage and land use reformation. Most of the participants felt that TOD would positively impact the environment of Riyadh City. For example, according to one participant, almost 18 L of fuel is burned daily due to increased reliance on private vehicles, so TOD would help in reducing this amount and therefore positively affect the Riyadh City environment. The majority of the participants felt that TOD could assist in reducing greenhouse gases by almost 10–24%. One of the participants stated that

TOD could reduce GHG by almost 35%, smog by 20%, and respiratory disease by 7.9% through proper transportation planning. The plan includes reducing the use of automobiles and increasing access to transportation.

4.1.3. Social Benefits

The majority of the participants stated that TOD would provide more destinations that could easily be assessed by bicycling and walking. According to one participant, TOD would influence physical activity in various ways, allowing people to walk to transit stops to access banks, restaurants, and grocery stores. Some of the participants stated that a reduction in the use of private vehicles would result in fewer accidents; however, it depends on the TOD design and how it is connected to the transportation system. Furthermore, it was stated by the interviewees that TOD would help in the reduction in mental health disorders. For example, one of the participants stated that driving on a congested road is mainly a great source of stress that the use of efficient TOD could reduce, as it would provide a more non-motorized mode of transportation.

Most of the participants in this study stated that TOD would positively impact the community by offering the benefits associated with this development. One of the participants indicated that TOD would provide more efficient and faster trips. Most participants indicated that TOD would benefit the community by reducing travel time and allowing people to adopt an active lifestyle. For example, according to one of the participants, TOD would reduce congestion and result in less parking costs. Another participant stated that TOD comprised good public transit, affordable housing, and cultural facilities that would help improve the livable communities' concept.

4.2. Questionnaire Findings

The quantitative data were collected using a questionnaire. Table 2 lists the findings of the TOD environmental benefits in Riyadh City from a general and personal perspective. From a personal perspective, the TOD benefit in terms of enhancing the city's aesthetic beauty and quality of city design obtained the highest mean score of 3.26. On the other hand, from a general perspective, the findings show that the maximum environmental benefit that could be achieved is that TOD would help in reducing traffic congestion, which would reduce noise and air pollution.

Table 2. Environmental Benefits of TOD.

	Findings	**Mean Higher Value**
General Perspective	TOD reduces traffic congestion and helps in improving air quality and noise pollution.	4.10
	TOD contributes to the land's more efficient use.	3.09
Personal Perspective	TOD enhances the quality of city design and improves aesthetic beauty.	3.26

Table 3 lists the findings of the TOD economic benefits in Riyadh City from a general and personal perspective. From a personal perspective, the economic benefit of TOD in terms of providing business opportunities obtained the highest score of 4.2. On the other hand, from a general perspective, the findings show that the maximum economic benefit that could be achieved is that TOD would encourage the use of public transport.

The findings of the TOD social benefits in Riyadh City from a general and personal perspective are shown in Table 4. From a personal perspective, the TOD benefit in providing services that enhance transit community mobility attained the highest mean score of 3.08. From a general point of view, the findings show that the maximum benefit that could be achieved is that TOD would provide options for mobility to older and young people who do not have cars.

Table 3. Economic Benefits of TOD.

	Findings	Mean Higher Value
General Perspective	TOD encourages the increased use of public transport.	4.08
Personal Perspective	TOD provides increased access to retail and local business.	4.20
	TOD increases property value.	3.81
	TOD would reduce the need for the construction of roads' cost associated with the long-term maintenance of roads.	3.95

Table 4. Social Benefits of TOD.

	Findings	Mean Higher Value
General Perspective	TOD provides various mobility options for older and young people who do not have their own car.	4.10
Personal Perspective	TOD has various service characteristics that increase automobile mobility.	3.08

5. Discussion

The present study used a mixed methodology to determine the benefits of transit-oriented development in Riyadh City, Saudi Arabia. The findings of this study show that TOD in Riyadh City would positively impact economic, environmental, and social aspects. One of the important findings of this study is that TOD would increase property value, which has been accepted at large. Previous studies have demonstrated that TOD has a positive effect on home values and land prices. For example, a study was carried out in Hong Kong, where it was found that TOD increased the housing price between 5–35% [20]. In another study carried out by Zhang, (2020) in Wuhan, China, it was found that properties that lay within 100 m of transit showed an increase of 17% in value, whereas those within 100–400 m showed an increase of 8% in their property value [11].

The study findings show that TOD would provide health benefits due to increased bicycling and walking. Considering this benefit, a study was conducted in Washington, which demonstrated that an increase in walking and cycling helps in reducing obesity. Many problems are related to obesity, such as diabetes, heart disease, cancer, and high blood pressure. Through TOD implementation, urban development and lower pollution can be achieved, which results in health benefits. The findings of this study demonstrate a relationship between physical activity and TOD. This finding is in line with studies conducted around the world [11,20]. A number of studies have demonstrated that transit-oriented development correlates with physical activity, irrespective of whether people live near TOD [21–23]. Studies have shown that transit travelers take walking trips more to access different services [21–23]. The findings of this study, along with the previous studies, clearly show that TOD increases walking activities. Considering these health benefits, further studies that involve medical records and direct measurements are needed to determine the actual health benefits associated with TOD implementation.

In the last few years, the reduction in the use of private vehicles has been the major focus of research because of the negative effect on the environment as well people's social life. The findings of this study show that implementing TOD would be beneficial in mitigating these issues. These findings are in line with previous studies, where it was demonstrated TOD would improve comfort and convenience by increasing access to activities and services, such as economic benefits, medical services, and financial savings, as well as enjoyment from being able to attend recreational activities [24,25].

In economic terms, TOD would bring economic benefits to Riyadh City. The findings of this study revealed that TOD would support the economy of Riyadh City by creating jobs. These findings align with previous studies, which stated that investing in TOD supports economic activities [26,27]. In one study, it was found that people living closer to the public transit service usually work more days every year as compared to those who lack access to this transit service [28]. Transit ridership would increase by implementing TOD, bringing additional revenues into the transit system and localities. This finding is in line with previous studies [26,27]. One study demonstrated that TOD would reduce the need for the construction of roads as well as the cost associated with the long-term maintenance of roads [28]. In another study, it was found that TOD would reduce the transportation cost of highways and externalities, such as infrastructure expenses, road maintenance, and fuel consumption costs [29]. The results of this study generated a general hope among most participants that implementing TOD would reduce household spending as they change their mode of transportation by switching to use public transport. Moreover, TOD would improve the accessibility of the city and save time and, at the same time, improve automobile mobility and reduce congestion costs.

From the point of view of environmental benefit, the findings show that TOD would help reduce environmental pollution by reducing the use of private vehicles. Most participants argued that TOD is needed to improve the city's air quality. Increasing awareness of TOD-related environmental considerations is important, as it would help discourage people from using private vehicles. These findings show that investing in TOD would provide the administration of Riyadh City with an opportunity to change their current system of transportation and to reduce the consumption of energy, which ultimately helps in creating an environmentally friendly city. Another study by Li et al. (2019) found that transit-oriented development is linked with environmental benefits [30]. It can be stated here that the more efficient the TOD in the Riyadh region, the more environmental benefits can be gained for the city, such as reducing pollution, enhancing air quality, and reducing congestion. Each of these benefits was found to be very important in the sustainable development of Riyadh City. Based on these findings, it is recommended that policymakers make strong policies so that the city can achieve environmental benefits and enhance the air quality with the implementation of TOD. The country is looking forward to diversifying its operations from an oil-oriented economy to a tourism-friendly economy and achieving the objective of environmental benefits.

Another important finding of this study is that social benefits are expected to increase for Riyadh City due to the implementation of TOD. Similar results were found in a study by Pendall et al. (2012), in which researchers linked transit-oriented development with social benefits and found that the more efficient the TOD, the more social benefits could be increased for the city [31]. Similar findings were found in another study by Higgins and Kanaroglou (2018), which found that TOD provides more efficient and faster trips by reducing travel time and allowing people to have an active lifestyle [32]. These findings imply that TOD is also important for the social well-being of people in Riyadh City, as it can help provide its citizens with all the fundamental rights. Based on the findings and previous studies conducted in the same domain, it is recommended that policymakers consider the concept of TOD in their region to bring more viability to their city, as this outcome would help achieve the 2030 vision.

6. Conclusions

This study aimed to evaluate the importance of transit-oriented development for the city of Riyadh, KSA. A mixed-study research was carried out. TOD is essential to strengthening Riyadh's social, environmental, and economic well-being. Regarding social benefits, TOD would help provide more mobility options for those without a car, the elderly, and the young. In terms of environmental benefits, TOD would improve the urban design quality and increase the urban landscape's aesthetic beauty. From an economic point of view, TOD would provide accessibility to alternative transport means that would increase the volume in terms of the mobility of people by a single system. This study's outcome provides useful insight into the benefits of TOD. There is a need to incorporate the combined action of various stakeholders to achieve the 2030 vision of a sustainable environment. To achieve significant transport development in Riyadh City, government authorities must show political commitment and form a comprehensive policy to integrate TOD into the urban transport system.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nieuwenhuijsen, M.J. New urban models for more sustainable, liveable and healthier cities post covid19; reducing air pollution, noise and heat island effects and increasing green space and physical activity. *Environ. Int.* **2021**, *157*, 106850. [CrossRef]
2. Qi, F.; Abu-Rumman, A.; Al Shraah, A.; Muda, I.; Huerta-Soto, R.; Hai Yen, T.T.; Abdul-Samad, Z.; Michel, M. Moving a step closer towards environmental sustainability in Asian countries: Focusing on real income, urbanization, transport infrastructure, and research and development. *Econ. Res.-Ekon. Istraživanja* **2022**, 1–20. [CrossRef]
3. Albuquerque, F.D.; Maraqa, M.A.; Chowdhury, R.; Mauga, T.; Alzard, M. Greenhouse gas emissions associated with road transport projects: Current status, benchmarking, and assessment tools. *Transp. Res. Procedia* **2020**, *48*, 2018–2030. [CrossRef]
4. Alajmi, R.G. Factors that impact greenhouse gas emissions in Saudi Arabia: Decomposition analysis using LMDI. *Energy Policy* **2021**, *156*, 112454. [CrossRef]
5. Khan, A.; Bibi, S.; Ardito, L.; Lyu, J.; Hayat, H.; Arif, A.M. Revisiting the dynamics of tourism, economic growth, and environmental pollutants in the emerging economies—Sustainable tourism policy implications. *Sustainability* **2020**, *12*, 2533. [CrossRef]
6. Papagiannakis, A.; Yiannakou, A. Do Citizens Understand the Benefits of Transit-Oriented Development? Exploring and Modeling Community Perceptions of a Metro Line under Construction in Thessaloniki, Greece. *Sustainability* **2022**, *14*, 7043. [CrossRef]
7. Abubakar, I.R.; Dano, U.L. Sustainable urban planning strategies for mitigating climate change in Saudi Arabia. *Environ. Dev. Sustain.* **2020**, *22*, 5129–5152. [CrossRef]
8. Rahman, M.M.; Hasan, M.A.; Shafiullah, M.; Rahman, M.S.; Arifuzzaman, M.; Islam, M.K.; Islam, M.M.; Rahman, S.M. A Critical, Temporal Analysis of Saudi Arabia's Initiatives for Greenhouse Gas Emissions Reduction in the Energy Sector. *Sustainability* **2022**, *14*, 12651. [CrossRef]
9. Oakil, A.T.; Anwar, A.M.; Alhussaini, A.; Al Hosain, N.; Muhsen, A.; Arora, A. Urban transport energy demand model for Riyadh: Methodology and a preliminary analysis. *Urban Plan. Transp. Res.* **2022**, *10*, 253–273. [CrossRef]
10. Carlton, I. Transit planners' transit-oriented development-related practices and theories. *J. Plan. Educ. Res.* **2019**, *39*, 508–519. [CrossRef]
11. Zhang, M. A Chinese approach to "transit oriented development" (TOD). In *Handbook on Transport and Urban Transformation in China*; Edward Elgar Publishing: Cheltenham, UK, 2020; pp. 116–130.
12. Knowles, R.D.; Ferbrache, F. Introduction to Transit Oriented Development and Sustainable Cities: Economics, Community and Methods. In *Transit Oriented Development and Sustainable Cities*; Edward Elgar Publishing: Cheltenham, UK, 2019; pp. 1–10.
13. Wogan, D.; Carey, E.; Cooke, D. *Policy Pathways to Meet Saudi Arabia's Contribution to the Paris Agreement*; King Abdullah Petroleum Studies and Research Center (KAPSARC): Riyadh, Saudi Arabia, 2019.
14. Hamieh, A.; Rowaihy, F.; Al-Juaied, M.; Abo-Khatwa, A.N.; Afifi, A.M.; Hoteit, H. Quantification and analysis of CO_2 footprint from industrial facilities in Saudi Arabia. *Energy Convers. Manag. X* **2022**, *16*, 100299. [CrossRef]
15. Noland, R.B.; Weiner, M.D.; DiPetrillo, S.; Kay, A.I. Attitudes towards transit-oriented development: Resident experiences and professional perspectives. *J. Transp. Geogr.* **2017**, *60*, 130–140. [CrossRef]
16. Papa, E.; Bertolini, L. Accessibility and transit-oriented development in European metropolitan areas. *J. Transp. Geogr.* **2015**, *47*, 70–83. [CrossRef]
17. Wey, W.-M.; Chiu, Y.-H. Assessing the walkability of pedestrian environment under the transit-oriented development. *Habitat Int.* **2013**, *38*, 106–118. [CrossRef]
18. Renne, J.L. Evaluating transit-oriented development using a sustainability framework: Lessons from Perth's network city. In *Planning Sustainable Communities: Diversity of Approaches and Implementation Challenges*; Tsenkova, S., Ed.; University of Calgary: Calgary, AB, Canada, 2009; pp. 115–148.
19. Zandiatashbar, A.; Hamidi, S.; Foster, N.; Park, K. The missing link between place and productivity? The impact of transit-oriented development on the knowledge and creative economy. *J. Plan. Educ. Res.* **2019**, *39*, 429–441. [CrossRef]
20. Cervero, R.; Murakami, J. Rail and property development in Hong Kong: Experiences and extensions. *Urban Stud.* **2009**, *46*, 2019–2043. [CrossRef]
21. Langlois, M.; Wasfi, R.A.; Ross, N.A.; El-Geneidy, A.M. Can transit-oriented developments help achieve the recommended weekly level of physical activity? *J. Transp. Health* **2016**, *3*, 181–190. [CrossRef]

22. Frank, L.D.; Fox, E.H.; Ulmer, J.M.; Chapman, J.E.; Braun, L.M. Quantifying the health benefits of transit-oriented development: Creation and application of the San Diego Public Health Assessment Model (SD-PHAM). *Transp. Policy* **2022**, *115*, 14–26. [CrossRef]
23. Durand, C.P.; Andalib, M.; Dunton, G.F.; Wolch, J.; Pentz, M.A. A systematic review of built environment factors related to physical activity and obesity risk: Implications for smart growth urban planning. *Obes. Rev.* **2011**, *12*, e173–e182. [CrossRef]
24. Litman, T. *Evaluating Public Transit Benefits and Costs*; Victoria Transport Policy Institute: Victoria, BC, Canada, 2015.
25. Litman, T. *Evaluating Public Transportation Health Benefits*; Victoria Transport Policy Institute: Victoria, BC, Canada, 2012.
26. Nesmachnow, S.; Hipogrosso, S. Transit oriented development analysis of Parque Rodó neighborhood, Montevideo, Uruguay. *World Dev. Sustain.* **2022**, *1*, 100017. [CrossRef]
27. Dutta, U.; Consortium, M.N.T.R. *Transit-Oriented Development (TOD) in Metro Detroit*; Mineta National Transit Research Consortium: Detroit, MI, USA, 2014.
28. Semeraro, L. *Approaches to Transit Oriented Development: How Existing Infrastructure Can Lend Itself to Implementation of Transit Oriented Development Techniques*; Columbia University: New York, NY, USA, 2018.
29. Mudigonda, S.; Ozbay, K.; Ozturk, O.; Iyer, S.; Noland, R.B. Quantifying transportation benefits of transit-oriented development in New Jersey. *Transp. Res. Rec.* **2014**, *2417*, 111–120. [CrossRef]
30. Li, Z.; Han, Z.; Xin, J.; Luo, X.; Su, S.; Weng, M. Transit oriented development among metro station areas in Shanghai, China: Variations, typology, optimization and implications for land use planning. *Land Use Policy* **2019**, *82*, 269–282. [CrossRef]
31. Pendall, R.; Gainsborough, J.; Lowe, K.; Nguyen, M. Bringing equity to transit-oriented development: Stations, systems, and regional resilience. *Build. Resilient Reg. Urban Reg. Policy Its Eff.* **2012**, *4*, 148–192.
32. Higgins, C.; Kanaroglou, P. Rapid transit, transit-oriented development, and the contextual sensitivity of land value uplift in Toronto. *Urban Stud.* **2018**, *55*, 2197–2225. [CrossRef]

sustainability

Article

Direct and Spillover Effects of Urban Land Expansion on Habitat Quality in Chengdu-Chongqing Urban Agglomeration

Sicheng Wang [1], Feng Lu [2] and Guoen Wei [3,*]

[1] College of Architecture and Urban Planning, Guizhou University, Guiyang 550025, China
[2] College of Architecture and Urban Planning, Chongqing University, Chongqing 400044, China
[3] College of Resources and Environment, Nanchang University, Nanchang 330031, China
* Correspondence: dg1927034@smail.nju.edu.cn

Abstract: Urban land expansion has dramatically changed the spatial distribution patterns and functional structure of habitats. Previous studies on the spatial externality effect of urban land expansion on the habitat quality of urban agglomerations are still insufficient. With the use of remote sensing and statistical data from 2000 to 2018, this study explored the evolutionary relationship between urban land expansion and habitat quality in the Chengdu-Chongqing urban agglomeration (CUA) using the bivariate local autocorrelation method and spatial Durbin model. Partial differential equation decomposition of the local and spatial spillover effects was implemented to investigate the marginal effects of the influencing factors. The highlights of the results are as follows: CUA's urban land increased by 2890.42 km^2 from 2000 to 2018, mainly caused by urban encroachment over farmland and grassland. New urban lands were situated primarily in the main urban districts of Chengdu and Chongqing; urban expansion intensity slowed to 7.64% in 2010–2018, declining by 53.95% from 2000 to 2010. The average habitat quality decreased to 0.905, and two "ring-shaped decline areas" were formed around the main urban areas of Chengdu and Chongqing. "Low-High" and "Low-Low" clusters were the main associations between urban land expansion and habitat quality changes. The impact of urban land expansion on local habitat quality changed from insignificant to negative, while its spatial spillover effects over adjacent areas have increased the negative environmental externalities to habitat quality in adjacent areas through spatial spillovers. Our findings provide evidence for urban agglomerations such as CUA that are still being cultivated to carry out cross-city joint protection strategies of habitat quality, also proving that habitat quality protection should be an integration of urban expansion regulation, natural adaptation and socioeconomic adjustment.

Keywords: urban land expansion; habitat quality; spatial spillover effect; spatial regression model; Chengdu-Chongqing urban agglomeration (CUA)

check for updates

Citation: Wang, S.; Lu, F.; Wei, G. Direct and Spillover Effects of Urban Land Expansion on Habitat Quality in Chengdu-Chongqing Urban Agglomeration. *Sustainability* **2022**, *14*, 14931. https://doi.org/10.3390/su142214931

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 16 September 2022
Accepted: 8 November 2022
Published: 11 November 2022

1. Introduction

Habitat quality can express regional biodiversity by assessing the extent of various urban habitat or vegetation types and the degradation levels of each of these types. To a certain extent, habitat quality can directly provide the functional strengths and weaknesses of regional ecosystems. The destruction and degradation of natural ecosystems are the primary causes of global biodiversity loss [1,2]. Due to rapid urban expansion, the fragmentation, degradation, and transformation of habitats impair vital ecosystem functions by reducing biomass and altering nutrient cycling and have profound impacts on global biodiversity [3,4]. By 2100, the global population is expected to grow to 10.88 billion (WPP2019), while the urban land area will increase by 1.8–5.9 times [5]. The explosive growth in urban population and land use would accelerate the decline in habitat quality, resulting in the loss of about 11 to 33 million hectares of natural habitat [6]. Such losses would be extremely detrimental to the implementation of the Future Earth strategy and the UN's Convention on Biological Diversity (CBD). China is one of the main countries with unprecedented urbanization and major conservation challenges. China is experiencing unprecedented urbanization, with 47.5% of the world's total new urban land area from 2000 to 2018, resulting in a sharp decline in biodiversity, habitat fragmentation and degradation, and numerous ecological problems [7]. These problems have become major obstacles for China in realizing the Post-2020 Global Biodiversity Framework set during the CBD COP15. To achieve China's Sustainable Development Goals (SDGs) by 2030, particularly SDG 11 (sustainable cities) and SDG 15 (sustainably managing habitat and halting biodiversity loss), a greater understanding of how urban land expansion affects the rate, scale, and spatial distribution of habitat quality loss is crucial.

Quantifying the spatio-temporal response of habitat quality on urban land expansion is of great value for optimal urban land management and ecological security protection. Previous studies have had significant differences in analyzing the impact of urban land expansion on habitat quality. Some focused on the negative effects of urban land expansion on habitat quality [8]. For example, Yang (2021) and Wang et al. (2022) found that urban land use expansion has changed the spatial pattern and functional elements of regional habitats, profoundly affecting the material and energy flows between habitat patches and playing a leading role in habitat quality decline [9,10]. They concluded that improving urban land use efficiency should be an important strategy to prevent and control habitat degradation. Some explored the regional heterogeneity of the impact of urban land expansion on habitat quality. For example, Feng et al. (2018) concluded that habitat degradation usually occurs in the functional expansion areas around cities and plain-mountain junctions, while the urban core's habitat quality usually improves gradually due to residential demand for habitat environment. Others have looked into how urban land expansion morphology affects habitat quality [11]. Dai et al. (2018) found that the fragmented spatial layout of built-up lands and increased morphological irregularities are the main negative factors contributing to habitat degradation in Changsha City. In addition, the changes in habitat quality in the future process of urban land expansion have also been investigated by some scholars [12]. Through simulation studies, Li et al. (2022) predicted that urban land expansion will accelerate the transformation and degradation of habitats and lead to biodiversity loss, especially among vertebrates [6]. Li et al. (2022), Gao et al. (2022), and Liu et al. (2022) also analyzed future changes in habitat quality in response to urban expansion from different regions [13–15]. However, given that previous studies have largely focused on cities, special terrain areas, or global perspectives, significant research gaps remain, particularly on the coercive effects of urban land expansion on habitat quality in urban agglomeration areas under the background of ecological integration construction.

More importantly, a large number of studies have analyzed the driving effects of urban land expansion on habitat quality using map visual analysis, coupled coordination model, and linear regression models, creating difficulties in quantifying the error effects of geospatial differences and limiting the practical value of the resulting estimates [16–18]. To address this issue, spatial regression models have been used in exploring the interactions between urban land expansion and ecological factors (e.g., carbon emissions and PM2.5 concentrations) [19]. For better analysis, driving effects are usually divided into local direct effects and spatial spillover effects [20]. For a long time, the spillover effect of urban land expansion on habitat quality in adjacent areas has been seriously underestimated, that is, the external impact of urban land expansion has been relatively under-considered in previous studies, which is obviously a lack of support for joint actions of regional habitat governance.

As highly integrated collections of cities, urban agglomerations have become important vessels of urbanization and urban land expansion in China. The Chengdu-Chongqing urban agglomeration (CUA) is one of the most important urban systems in China's existing urban agglomeration hierarchy, the others being the urban agglomerations of Beijing-Tianjin-Hebei, Guangdong-Hong Kong-Macao Great Bay Area, Yangtze River Delta, and the middle reaches of the Yangtze River (Figure 1). According to the China Statistical Yearbook, CUA is one of the core regions of China's economy, population, and technology, accounting for 6.20% of GDP and 6.81% of the population in 2018. The region, having rich natural resources and abundant wildlife (e.g., cliff cypress, silver fir, ginkgo and golden monkey), is an important treasure trove of biodiversity in China. However, the rapid development of the industrial economy and population agglomerations have led to accelerated urban expansion and the conversion of grasslands and forests into construction lands, putting significant pressure on wildlife habitats and biodiversity [21]. Additionally, while the "Ecological Environment Protection Plan of Chengdu-Chongqing Twin Cities Economic Circle" by the regional government has been recently established to strengthen biodiversity investigation and monitoring in ecologically sensitive areas, understanding the impact mechanisms of urban land expansion on the habitat quality in the CUA has been limited.

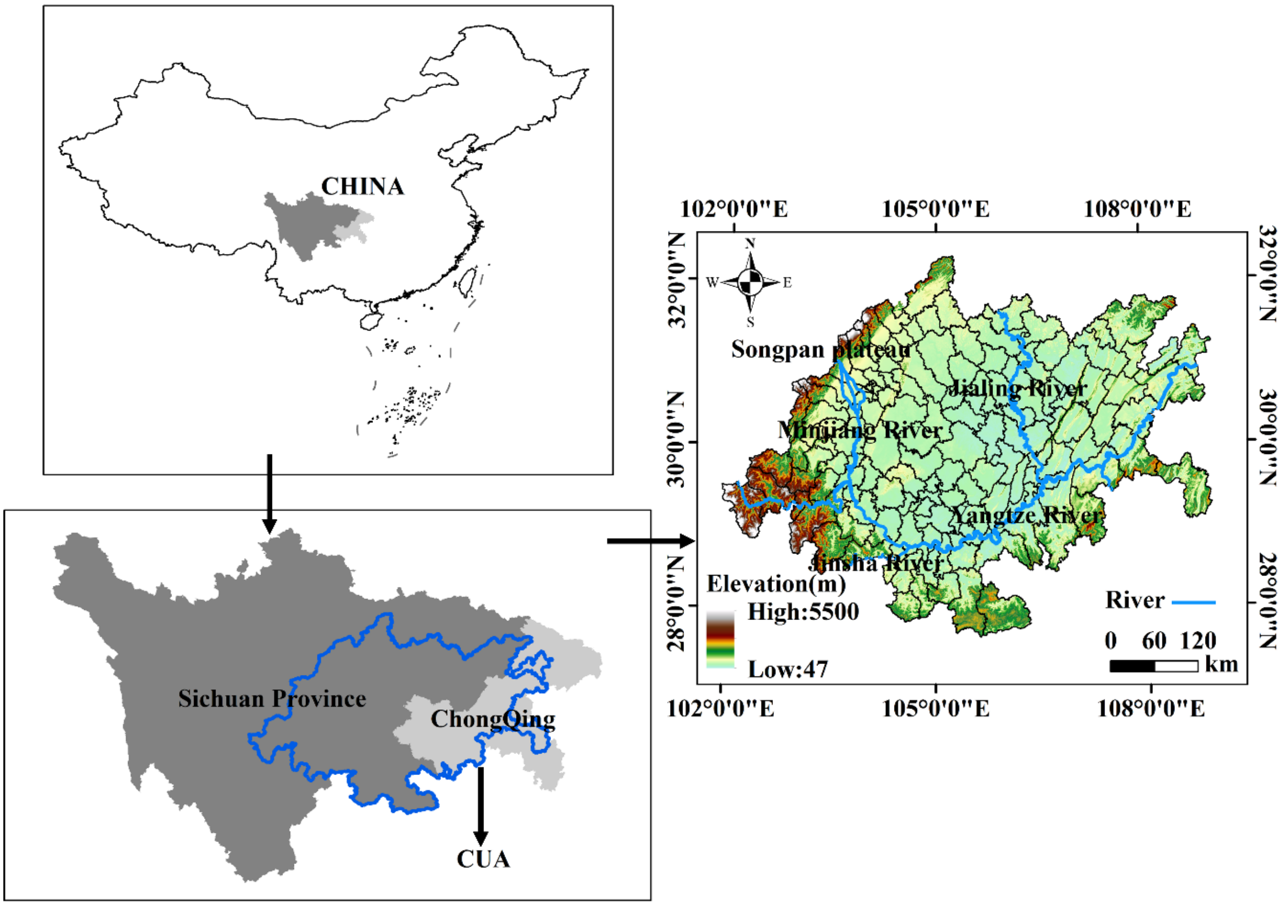


Figure 1. The geographical location of CUA.

In summary, previous studies have made substantial contributions in exploring the effects of urban land expansion on habitat quality, providing useful references for regional ecological security protection and urban land management. However, the discussion on urban land expansion and habitat quality in the context of rapid urbanization is far from settled due to the following reasons: (1) the impact mechanism of urban land expansion on habitat quality has largely been overlooked at the urban agglomeration level; (2) the spatial externality of the urban land expansion on habitat quality has long been neglected; (3) to our knowledge, there has been no study exploring the impact mechanism of urban land expansion on habitat protection in CUA. To fill these knowledge gaps, this study explored the differentiated mechanisms of the local direct and spatial spillover effects of urban land expansion on habitat quality in the CUA. Google Earth Engine (GEE) geographic cloud platform was used to obtain the urban land expansion data from 2000 to 2018, and the InVEST-Habitat Quality model was applied to describe the spatio-temporal distribution and trend changes in habitat quality. The bivariate local spatial autocorrelation method was used to analyze the spatial correlation between urban land expansion and habitat quality, while the Spatial Durbin model evaluated the spatial externalities of the impact of urban land expansion on habitat quality. It is worth mentioning that these research paths were realized using administrative and geographical grid levels to increase the study's objectivity.

2. Study Area, Methods and Data

2.1. Study Area

CUA has emerged as a critical platform for China's western development strategy, providing vital support in the development of the Belt and Road Initiative [22]. CUA is centered in Chongqing and Chengdu and includes Zigong, Luzhou, Deyang, Mianyang, Suining, Neijiang, Leshan, Nanchong, Meishan, Yibin, Guang'an, Dazhou, Ya'an, and Ziyang. To ensure study refinement and adequacy of the sample size in the regression analysis, the regression analysis of our study was conducted at the county level, resulting in a total of 142 samples (Figure 1).

2.2. Methods

The research methods used are shown in Figure 2, and they were used to investigate the effects of urban expansion intensity on habitat quality in the CUA.

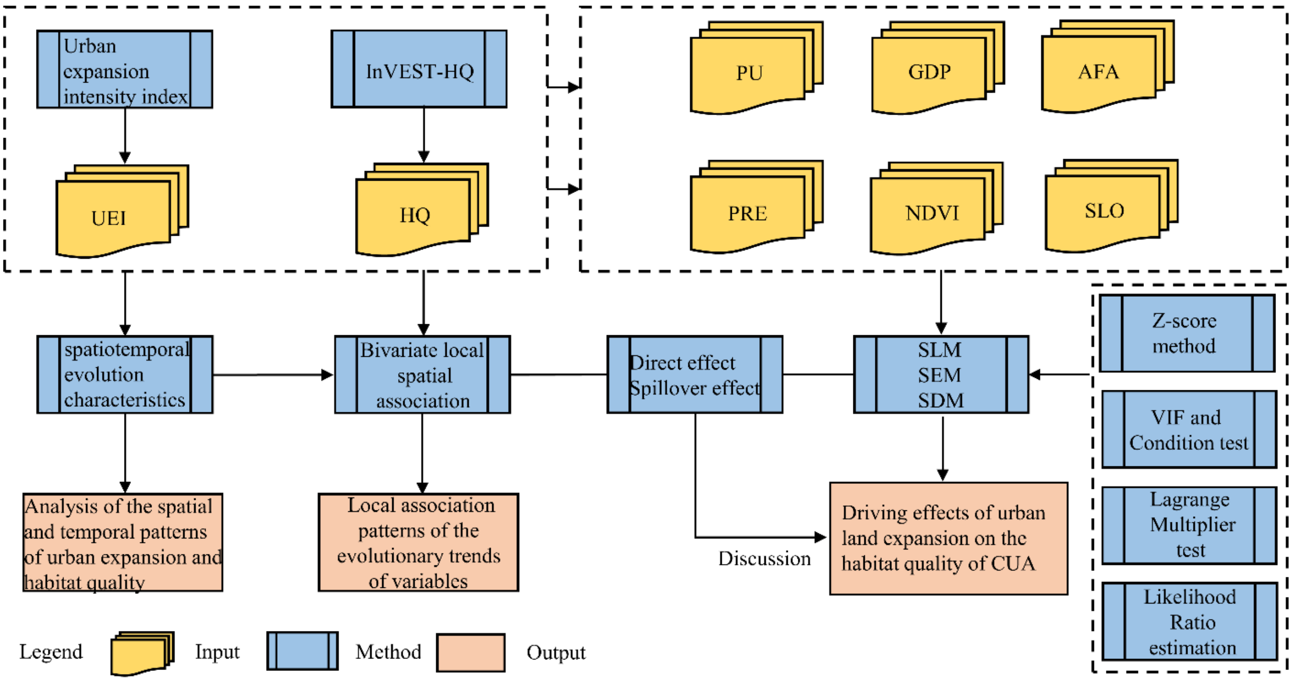


Figure 2. Study Flowchart (the original meaning of abbreviations, see Abbreviation for detail).

2.2.1. Urban Land Expansion Intensity Index

The urban land expansion intensity index is used to characterize the expansion degree of urban areas in the CUA. This index can measure the state of expansion of a region at varying stages and compare the development intensities of different regions at same stages [23]. The calculation formula is as follows:

$$UEIn = \frac{A_n^{t2} - A_n^{t1}}{A_n^{t1} \times \Delta t} \times 100\% \quad (1)$$

where UEI_n is the urban land expansion intensity index of the nth study unit; A_n^{t1} and A_n^{t2} are the urban land areas at nodes t_1 and t_2, respectively; Δt is the interval year from t_1 to t_2.

2.2.2. InVEST-Habitat Quality Model

Integrated Valuation of Ecosystem Services and Trade-offs (InVEST) is a model system developed by the U.S. Natural Capital Project team for valuing ecosystem services and their economic value to support ecosystem management and decision-making [24,25]. Among them, the Habitat Quality (HQ) module evaluates the biodiversity of the study area through the level of habitat degradation and suitability, reflecting the potential ability of the ecosystem to provide living and breeding conditions for species [26]. Habitat quality measurement includes two aspects: habitat degradation and habitat suitability. The former refers to the disturbance intensity of the threat source to the habitat, while the latter refers to habitat suitability.

The HQ module measures the habitat degradation degree with various parameters (e.g., threat source, threat source sensitivity, and distance between habitat and threat source) using the following equations:

$$D_{xj} = \sum_{r=1}^{R} \sum_{y=1}^{Y_r} r_y \left(\frac{\omega_r}{\sum_{r=1}^{R} \omega_r} \right) \times i_{xy} \beta_x S_{jr} \tag{2}$$

$$i_{rxy} = 1 - \left(\frac{d_{xy}}{d_{r\max}} \right) \text{ (if it is linear decay)} \tag{3}$$

$$i_{rxy} = \exp\left(\frac{-2.99 d_{xy}}{d_{r\max}} \right) \text{(if it is exponential decay)} \tag{4}$$

where D_{xj} is the degree of habitat degradation; R is the number of threat sources; Y_r is the grid number of threat sources; ω_r is the weight of threat source; r_y is the stress value of grid y; i_{rxy} is the stress level of grid y to grid x; β_x is the accessibility of threat source to grid x; S_{jr} is the sensitivity of the habitat type j to the threat source r; d_{xy} is the Euclidean distance of the habitat to the threat source; d_{rmax} is the maximum disturbance radius of the threat source r to the habitat.

The degree of habitat degradation (D_{xj}) and the habitat suitability are then used for the comprehensive evaluation of habitat quality, see the equation:

$$Q_{xj} = H_j \left[1 - \left(\frac{D_{xj}^z}{D_{xj} + k^z} \right) \right] \tag{5}$$

where Q_{xj} is the habitat quality index, the larger the value, the better the regional habitat quality and biodiversity service function; H_j is the habitat suitability of land-use type j; k is the half-saturation constant, which is half of the maximum degradation degree; z is the normalization constant, which is generally taken as 2.5 with reference to the previous studies by scholars [26].

As for the threat source, it is generally believed that the higher the degree of human utilization, the higher the threat to the habitat, and the greater the impact on the surrounding habitat biodiversity and ecosystem. Therefore, in this study, the areas of intensive human activities, such as farmland and urban land, were considered the main threat source areas. Bare lands, such as saline soils, barren grounds, and sandy lands, were also set as habitat threat factors for their poor ecological conditions and erosive effects on external habitats and the stability of surrounding ecosystems [27]. It is generally believed that the more pristine and complex the ecosystem, the higher the habitat suitability. Land-use types such as forest land, grasslands, and water areas were considered habitats in the study. The maximum impact distance and weight of each threat factor and the suitability and sensitivity of each habitat type were set according to the InVEST operation manual and the parameter settings of previous habitat quality studies in different regions of China (see Tables 1 and 2).

Table 1. Threat source and their maximum impact distance and weight.

Threat Source	The Maximum Influence Distance/km	Weight	Space Decline Type
Farmland	4	0.3	Exponential decay
Urban land	10	1	Exponential decay
Bare land	3	0.1	Exponential decay

Table 2. Habitat suitability of land-use types and the sensitivity to threat factors.

Land Use Type	Habitat Suitability	Threats		
		Farmland	Urban Land	Bare Land
Farmland	1	0.5	0.65	1
Forest land	1	0.6	0.8	1
Grassland	0.75	0.5	0.55	1
Water area	1	0.6	0.75	1
Urban land	0	0	0	0
Bare land	0	0	0	0

2.2.3. Spatial Driving Effect Analysis

Bivariate spatial autocorrelation methods and spatial regression models were used to analyze the spatial association patterns and driving effects of urban expansion and habitat quality changes, respectively. Among them, bivariate spatial autocorrelation is an extension of traditional spatial autocorrelation analysis, used to test the spatially coupled correlation between two variables [28]. In different regions, the response of habitat quality to urban land expansion generally has local non-stationarity, which can be reflected and identified using the bivariate LISA map. Based on the Multivariate LISA tool of GeoDa 1.4.6, a bivariate local spatial autocorrelation method was employed to quantify the local variability of the association between urban land expansion and habitat quality in the CUA. The calculation formula for the bivariate Moran's I is:

$$I_{kl}^{i} = \frac{x_k^i - \overline{x}_k}{\sigma_k} \cdot \sum_{j=l}^{n} \left(W_{ij} \frac{x_l^i - \overline{x}_l}{\sigma_l} \right) \tag{6}$$

where W_{ij} is the spatial weight matrix; x_k^i is the observation value k of study unit i; x_l^i is the observation value l of study unit j; σ_k and σ_l are the variances of x_k and x_l. The value range of I is between $[-1, 1]$. Values greater than 0 indicate positive correlations, in which similar variables tend to be clustered in space. Values less than 0 suggest negative correlations, in which similar variables tend to be discrete.

The spatial regression model can be used to further illustrate the local direct and spatial spillover effects of independent variables. There are mainly three common spatial regression models. The Spatial Durbin Model (SDM) integrates the ability of two methods (Spatial Error Model (SEM) and Spatial Lag Model (SLM)) to quantify the exogenous and endogenous interaction effects of variables and classifies the driving effects of urban land expansion on habitat quality into local direct effects and spatial spillover effects [29]. The SDM formula is as follows:

$$Yit = \rho WYit + \beta Xit + \theta WXit + \varepsilon it \tag{7}$$

where Y_{it} is the explained variable of region i in period t, expressed as increments in habitat quality; X_{it} is the explanatory variable of region i in period t, including urban land expansion and control variables; ρ, β, and θ are the parameters to be estimated; ε is the random disturbance term of the normal distribution; W is the spatial weight matrix; WY is the spatial lag dependent variable; WX is the spatial lag independent variable. Based on previous studies practice and the comparison of model operation results, the queen contiguity method and Euclidean distance were used to construct a spatial weighting matrix that reflects the spatial structure and location relationship of the data [27].

To improve the scientificity of the model operation, the following preprocessing steps were performed on the variable data. First, normalization of variable data (Z-score method) was performed while maintaining a normal distribution to improve the comparability of regression coefficients. Second, Variance Inflation Factor (VIF) and Condition index were used to test for possible multicollinearity among explanatory variables. Third, tests were conducted to evaluate the necessity of spatial regression and the rationality of model selection before constructing the SDM models of habitat quality drivers [21,29]. The Lagrangian Multiplier (LM) was used to test the necessity of incorporating spatial effects into the regression model. The Likelihood Ratio estimation (LR) was used to assess whether the SDM model can be reduced to SLM or SEM, that is, whether SDM can integrate the measure advantages of SLM and SEM. In analyzing the driving effects, the SDM model was used to divide the spatial regression effects into total effects, local direct effects, and spatial spillover effects using partial differential equations (P.D.E), focusing on the spatial characteristics of the driving effects of urban land expansion on habitat quality [30].

2.3. Data Source

Details of the variables used in this paper can be found in Table 3. The required data were obtained from the following sources:

Table 3. Variable category.

Variable Category	Variable	Abbreviation	Unit
Socioeconomic	Urban expansion intensity index	*UEI*	%
	Population urbanization level	*PU*	%
	GDP	*GDP*	yuan
	Agricultural fertilizer application	*AFA*	t
Natural	Habitat quality	*HQ*	-
	Slope	*SLO*	°
	Average annual precipitation	*PRE*	mm
	Normalized Difference Vegetation Index	*NDVI*	-

Land cover data. Land cover was determined using the global land cover dataset for 1992–2020 from the European Space Agency (ESA-CCI) (http://due.esrin.esa.int/globcover/) (accessed on 20 June 2021). With a 300 m resolution, the dataset has an overall accuracy of 75.38%, among which the user accuracy of urban land is 88%, which has a good interpretation effect on the current situation of regional land use. Using the land-use classification system of the Intergovernmental Panel on Climate Change (IPCC) and previous studies, the dataset was adjusted on the GEE geographic cloud platform and merged into six general categories: farmland, forest land, grassland, water area, urban land, and bare land [31].

Socio-economic data. In addition to urban land expansion, regional habitat quality has been shown to be associated with socio-economic factors such as economic development, urban population size, and food production [25,32,33]. Accordingly, on the basis of considering the prevalence of variables and existing research, GDP (*GDP*), population urbanization level (*PU*) and agricultural fertilizer application (*AFA*) were used to indicate the error effect of regional economic level, urban population size and food production [34]. These socio-economic data were acquired mainly from the "Statistical Yearbook of Sichuan Province" and the "Statistical Yearbook of Chongqing city" at the county and district levels.

Natural control variables. Indicators such as slope, precipitation, and Normalized Difference Vegetation Index (NDVI) have also been shown to have important effects on habitat quality [33,35]. Therefore, we included natural factors such as slope (*SLO*), precipitation (*PRE*), and NDVI (*NDVI*) as control variables in the model. The slope data had a 30 m spatial resolution and were derived from the GDEMV2 DEM digital elevation product of the Computer Network Information Center of the Chinese Academy of Sciences (http://www.gscloud.cn) (accessed on 15 January 2022). Spatial distribution data of precipitation were obtained from the spatial dataset of meteorological conditions provided by the National Earth System Science Data Center (https://www.resdc.cn/) (accessed on 20 January 2022). The NDVI dataset was derived from satellite remote sensing such as SPOT/VEGETATION and MODIS, and the vegetation cover distribution data were obtained through mosaic and projection transformation (https://www.resdc.cn/data.aspx?DATAID=343) (accessed on 30 January 2022).

3. Results

3.1. Spatio-Temporal Evolution Characteristics of Urban Land Expansion

3.1.1. Changes in Land Use Structure

Using the statistical analysis in ArcGIS (see Figure 3), farmland and forest land were the main land-use types in the CUA from 2000 to 2018, accounting for more than 95% of the region. Urban lands can be found in the plains and riverside areas of the middle of the urban agglomeration, concentrated mainly in the central regions of Chengdu and Chongqing. Table 4 summarizes the area transition matrix for the different land-use types at different periods. From 2000 to 2018, urban land expanded the most, increasing by 2890.42 km^2, equivalent to 170.82%. Forest land increased by 752.78 km^2, growing by 1.92%, while farmlands declined by 2.88% (4074.74 km^2). The increase in urban lands was caused mainly by encroachment into farmlands and grasslands and was particularly pronounced from 2010 to 2018 when 1511.36 km^2 of farmland was converted into urban land, accounting for 1.08% of the total area of farmland in 2005.

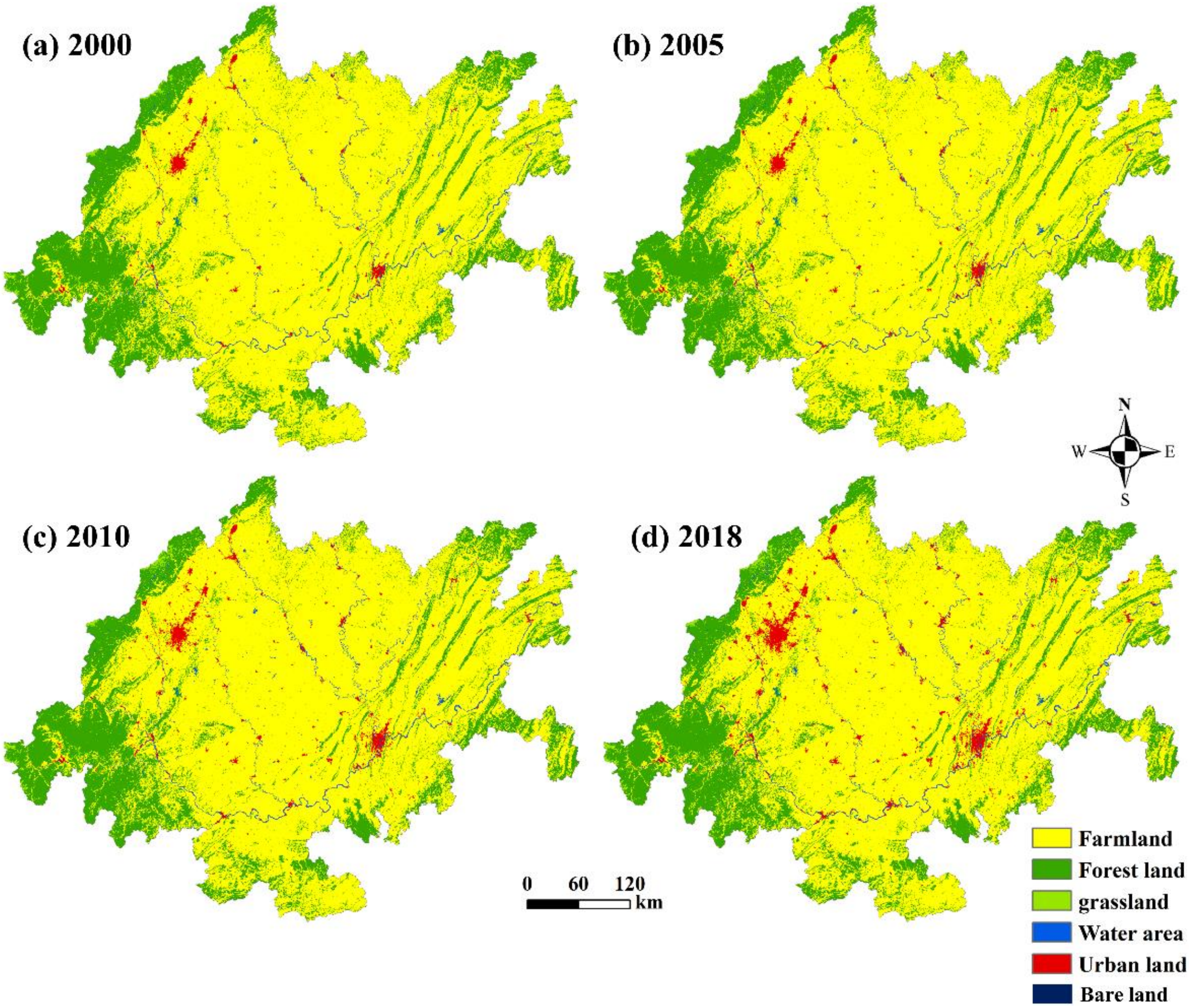


Figure 3. Distribution and evolution of land-use types in CUA from 2000 to 2018.

Table 4. The transition matrix of land-use types (km^2).

Year	Land Use Types	Farmland	Forest Land	Grassland	Water Area	Urban Land	Bare Land
2000–2005	Farmland	140,507.95	461.76	0.00	0.00	360.65	0.00
	Forest land	169.78	38,876.04	19.36	1.82	1.24	0.00
	Grassland	1.74	6.29	579.74	0.00	20.44	0.00
	Water area	0.00	0.00	0.00	2167.31	7.12	0.00
	Urban land	0.00	0.00	0.00	0.00	1691.90	0.00
	Bare land	0.00	0.00	0.00	0.00	0.00	0.00
2005–2010	Farmland	139,770.18	21.93	6.21	5.05	876.11	0.00
	Forest land	146.53	39,068.57	105.33	19.94	3.72	0.00
	Grassland	0.08	0.17	572.13	0.08	26.64	0.00
	Water area	0.00	0.00	0.00	2152.26	16.88	0.00
	Urban land	0.00	0.00	0.00	0.00	2081.35	0.00
	Bare land	0.00	0.00	0.00	0.00	0.00	0.00
2010–2018	Farmland	137,546.19	842.93	1.57	14.73	1511.37	0.00
	Forest land	78.52	38,972.76	7.53	15.47	16.38	0.00
	Grassland	0.08	5.30	651.56	2.23	24.49	0.00
	Water area	0.00	0.00	0.00	2152.09	25.23	0.00
	Urban land	0.00	0.00	0.00	0.00	3004.70	0.00
	Bare land	0.00	0.00	0.00	0.00	0.00	0.00

3.1.2. Spatio-Temporal Distribution Characteristics of Urban Land Expansion Intensity

Figure 4 shows the evolution of urban land expansion in the CUA. During the study period, the core areas of urban land expansion were in the main urban areas of Chengdu and Chongqing, while urban land expansion in small and medium-sized cities, such as Zigong, Luzhou, and Deyang, was relatively limited. Urban land expansion had a "Core-Periphery" gradient expansion pattern, and the extent of urban land expansion was most prominent from 2010 to 2018 compared to previous periods.

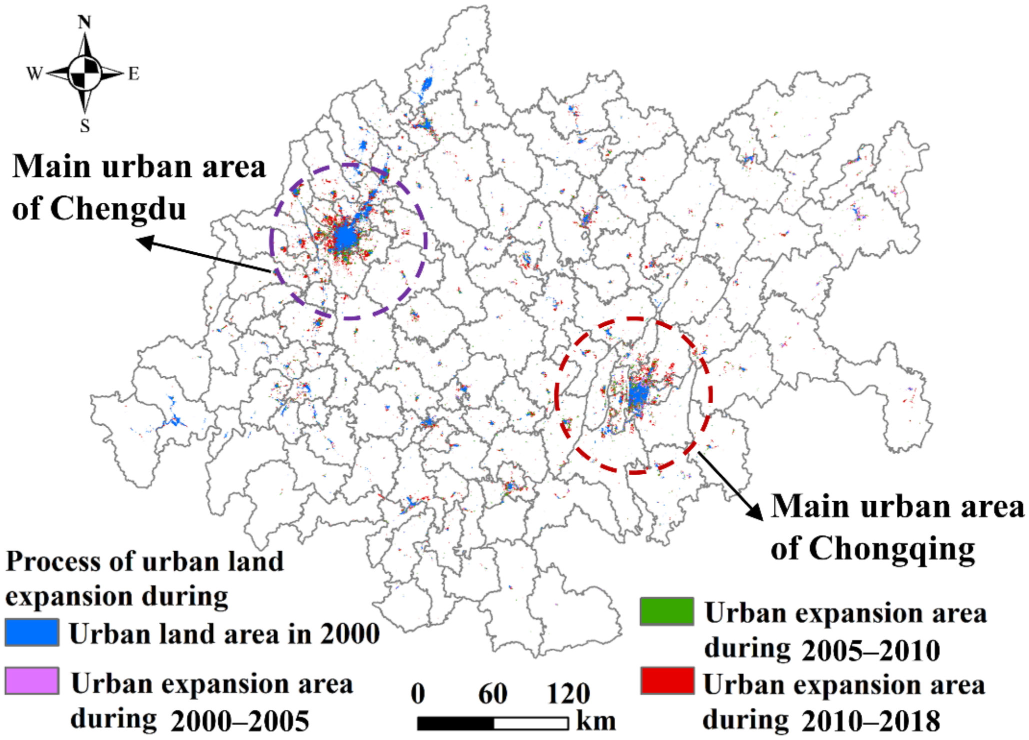


Figure 4. The process of urban land expansion from 2000 to 2018.

The overall urban land expansion of the CUA showed a slowing trend. Compared with 2000–2010, the average UEI index of the urban agglomeration dropped to 7.64% in 2010–2018, decreasing by 53.95% (Figure 5). The spatial distribution pattern of the UEI changed significantly during the 19-year research period. We classified the UEI index of the CUA region into five levels from "V-I" based on the natural breakpoint classification with nodes of 5%, 15%, 50% and 100%. From 2000 to 2010, 34.51% and 38.73% of county units were in levels IV and III in the UEI index. Level IV areas were concentrated in Leshan, Meishan, Neijiang, and southern Chongqing counties, while the Level III areas were situated mainly in Ziyang, Luzhou, and western Chongqing counties. From 2010 to 2018, the urban land expansion rates in many counties generally slowed down, possibly due to restrictive policies on regional land development. Cities classified as Level IV became dominant, accounting for 53.52% of all county-level units, followed by Class V, which comprised 40.85%. In addition, the UEI index distribution also significantly changed between 2000–2005 and 2005–2010, evolving from a configuration of "low-speed in the west and high-speed in the east" into a scattered distribution of low-speed expansion areas (V-level).

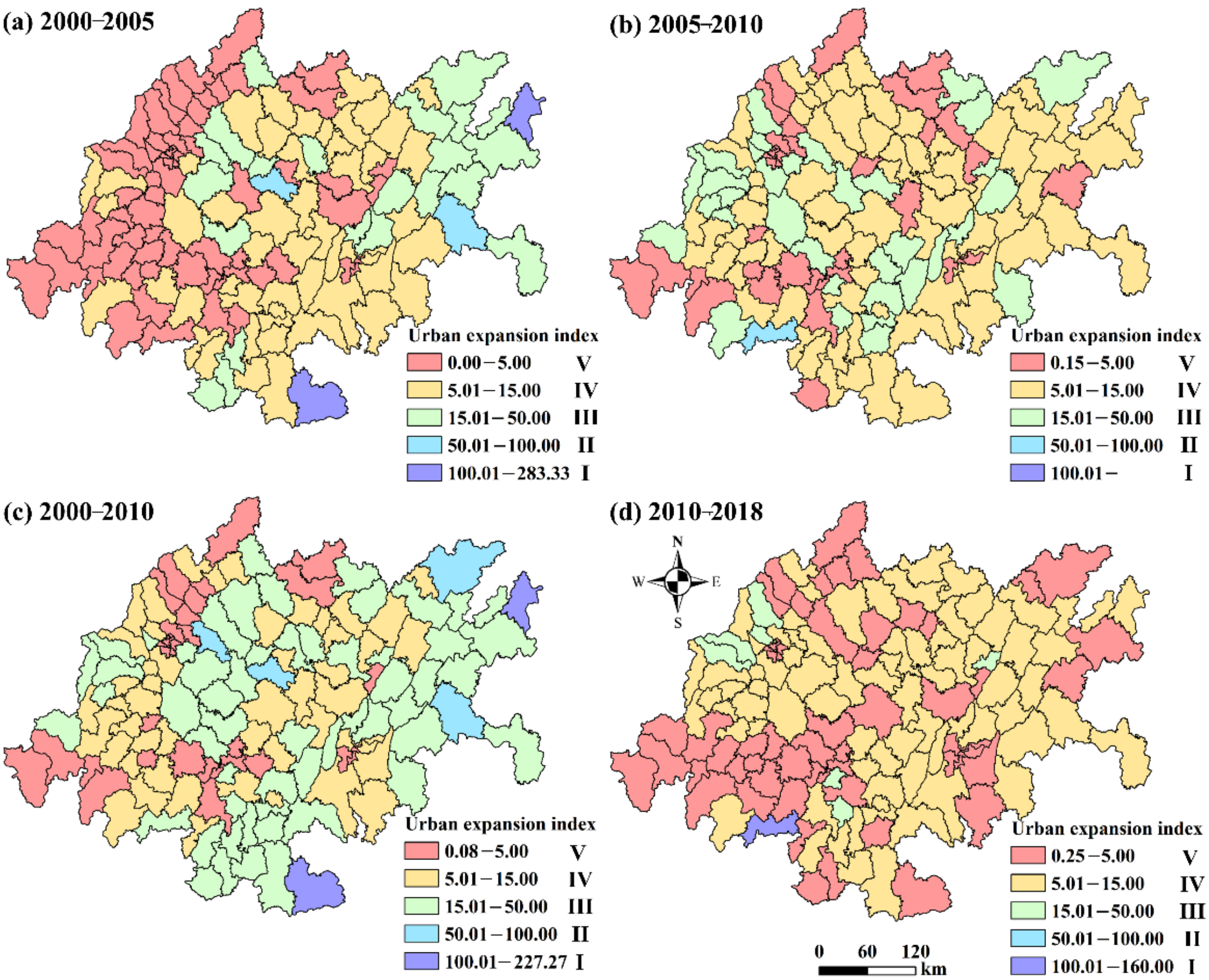


Figure 5. Spatio-temporal distribution pattern of UEI from 2000 to 2018.

3.2. Spatio-Temporal Evolution Characteristics of Habitat Quality

Figure 6 presents the evolution of the spatio-temporal distribution of habitat quality in the CUA. Habitat quality deteriorated at the administrative level for the given research period, with the average HQ index decreasing by 4.03%, from 0.943 in 2000 to 0.905 in

2018. Spatially, the habitat quality in the CUA exhibited pronounced regional differences. High-value areas were clustered in the Jiajin Mountain, Daliang Mountain, and Qionglai Mountain in the southwest, including the cities of Ya'an, Leshan, and Yibin, which have relatively low economic development levels. The low-value areas in Chengdu and Chongqing expanded considerably during the study period, their central urban districts of Jinjiang, Chenghua, Wuhou, and Yuzhong had the lowest habitat quality in the urban agglomeration.

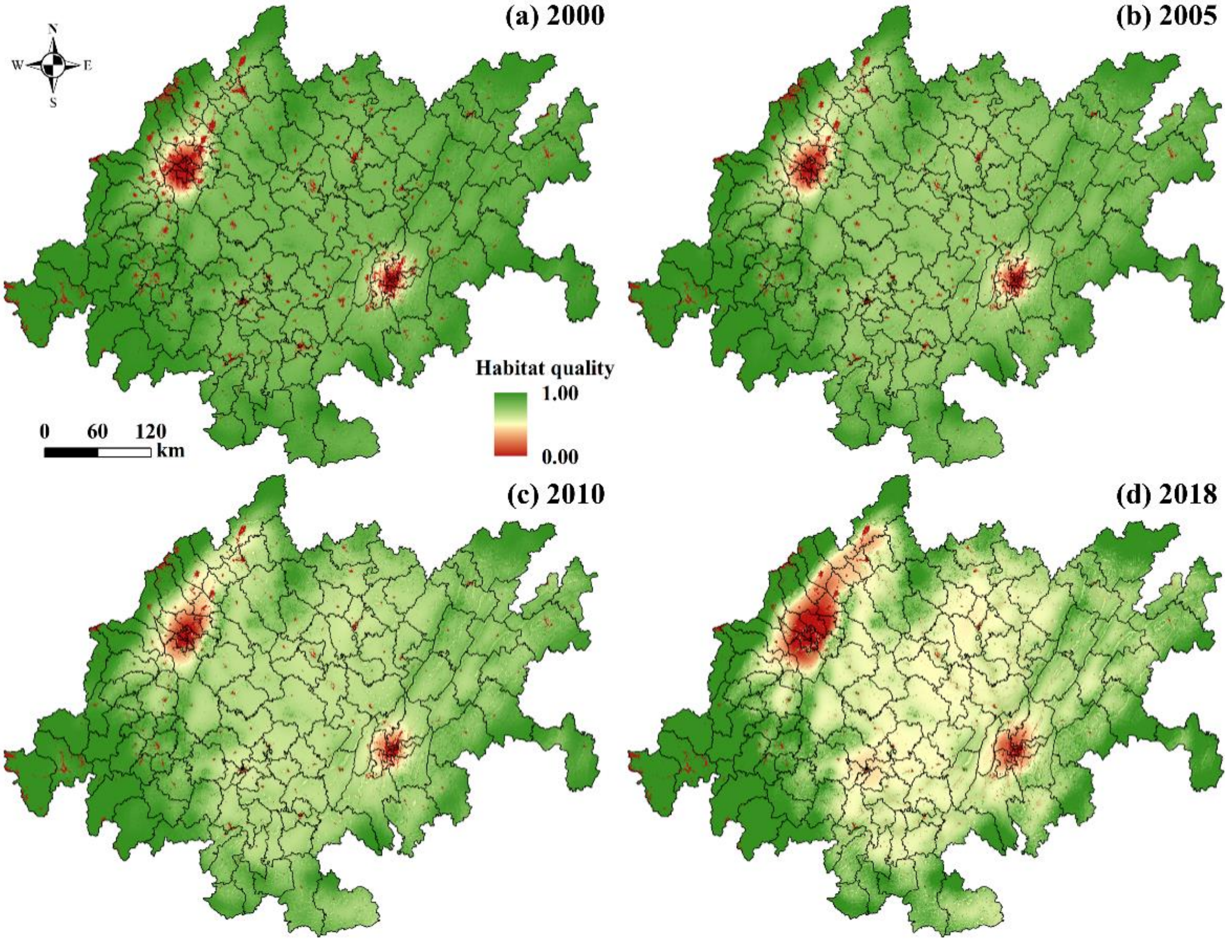


Figure 6. Spatio-temporal distribution pattern of habitat quality from 2000 to 2018.

We then analyzed the trend in habitat quality by subtracting the habitat quality at different periods (see Figure 7). The most significant decline in habitat quality was due to the peripheral expansion of the main metropolitan centers; in particular, two areas with considerable reduction in habitat quality developed around the main urban areas of Chengdu and Chongqing in a ring-shaped formation. Habitat quality in the marginal areas (e.g., southwestern and southeastern mountainous areas) significantly improved, particularly in the administrative regions of Mabian County, Ebian County, Junlian County, Xuyong County, and Qianjiang District. In addition, the area of "ring-shaped decline areas" with declining HQ values and habitat growth areas both had slow growths, increasing the structural complexity of the habitat system in the CUA.

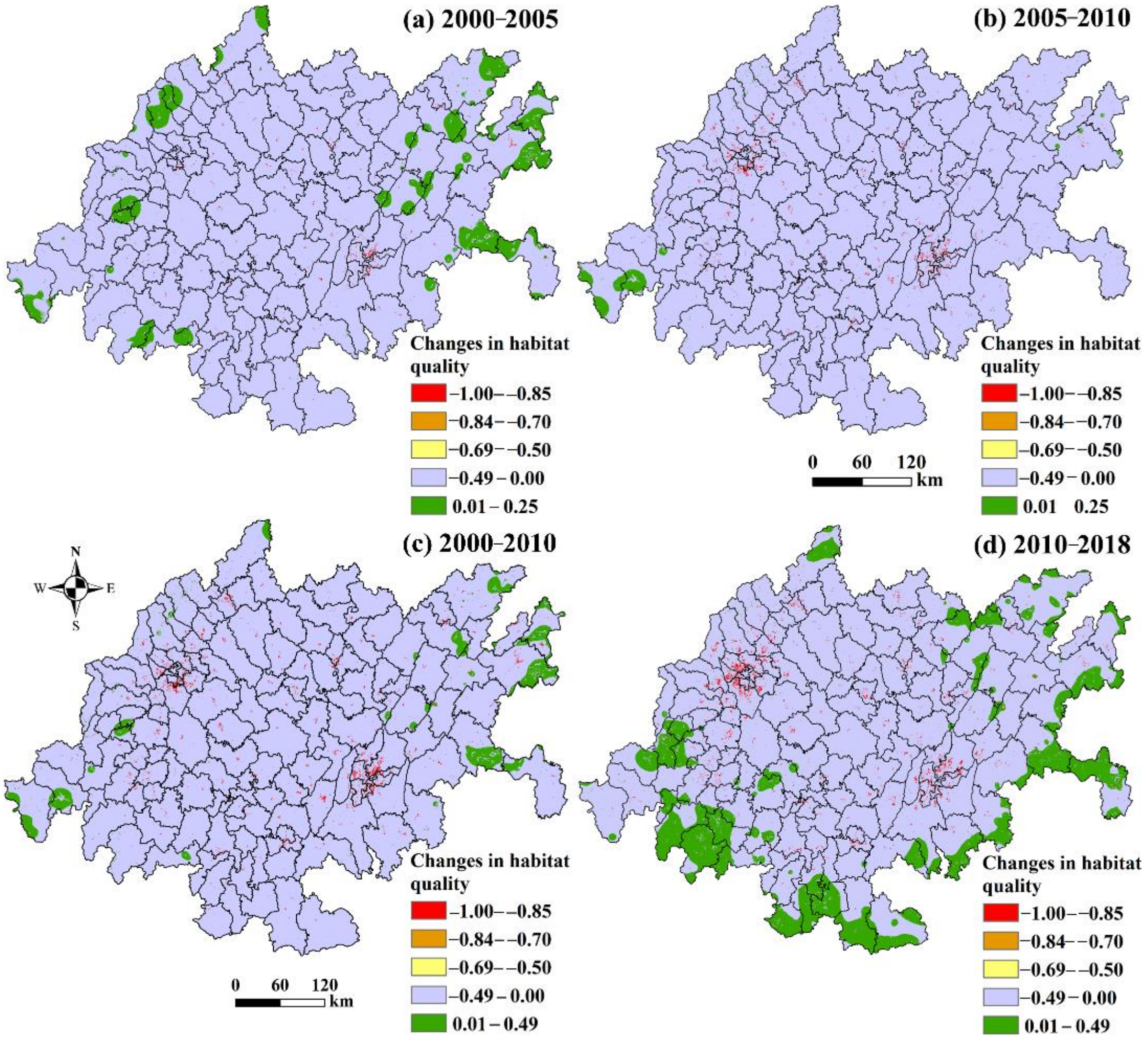


Figure 7. Spatial distribution changes in habitat quality from 2000 to 2018.

3.3. Driving Impacts of Urban Land Expansion on Habitat Quality

3.3.1. Bivariate Spatial Autocorrelation between Urban Land Expansion and Habitat Quality

The spatial analysis tool of the GeoDa platform was used to calculate the global spatial autocorrelation index Moran's I for urban land expansion and habitat quality changes. The Moran's I index for urban land expansion and habitat quality changes was 0.120 for 2000–2010 and 0.019 for 2010–2018. Since Moran's I indexes were low and exhibited a downward trend, the values cannot strongly prove a positive correlation between urban land expansion and habitat quality. This means that further analysis of the local variability of association is needed through bivariate local autocorrelation.

The investigation emphasizes that "Low-High" negative correlation and "Low-Low" positive correlation are the main agglomeration types associated with urban land expansion and habitat quality changes in the CUA (Figure 8). From 2000 to 2010, the negative correlation cluster regions accounted for 6.34% of the total area were formed by southwestern counties, such as Shimian County, Hanyuan County, Xuzhou District. Indicating the urban development intensity in these areas was relatively weak, while the habitat quality and regional ecological conservation efforts were great. "Low-Low" positive correlation cluster regions accounted for 11.27% of the total area and were consistent with the old urban areas of Chengdu and Chongqing, such as Chenghua District, Jiangjin District. This may be related to the early infrastructure construction that destroyed the natural habitat of the old urban areas, but the current urban construction space tends to be saturated. "High-High" positive correlation cluster regions accounted for 4.93% of the total area were formed by some southwestern mountainous counties, such as Xinjing County and Gulin County, which may be associated with the high-quality ecological base of these areas but the rapid acceleration of urban construction in recent years. As the new urban area of Chengdu and Chongqing, Longquanyi District, Beibei District, Yubei District, and Bishan District undertake the function of relieving the pressure of land demand in the old urban area, forming "high-low" negative correlation cluster.

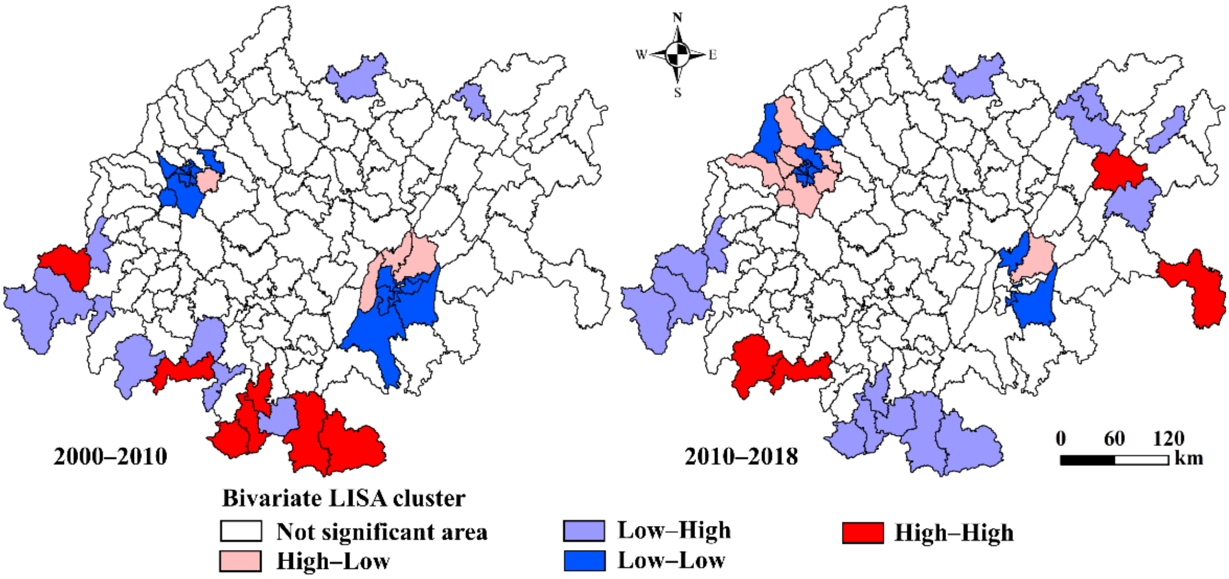


Figure 8. The spatial distribution of bivariate LISA between urban expansion and habitat quality in 2000–2018.

The number of counties with cluster types "Low-High" or "High-Low", where urban land expansion is negatively correlated with habitat quality, increased from 9.15% in 2000–2010 to 16.90% in 2010–2018. This means that the adversarial relationship between urban land expansion and habitat quality has become more prominent and that reconciling the relationship between urban development and ecological protection is extremely important. With the growth in land-use demand in old cities due to rapid population agglomeration, new cities around Chengdu and Chongqing (e.g., Chongzhou City, Pengzhou City, and Pidu District) have become new "High-Low" negative correlation cluster areas, accounting for 6.34% of the total area. "Low-high" cluster areas increased to 10.56%, and the newly added areas included northern and southern counties in the CUA, such as Kaizhou District, Xuyong County, and Gulin County. In comparison, correlated clusters of "Low-Low" and "High-High", where urban land expansion and habitat quality have a positive correlation, shrank to 8.45% and 2.82%, respectively. The results suggest that the phenomenon of deviation from the fundamental conflict situation of "urban land expansion-habitat quality" in the global autocorrelation analysis is gradually disappearing.

3.3.2. Further Analysis of the Driving Effect of Urban Land Expansion on Habitat Quality

Based on the data preprocessing results (i.e., logarithmization of variables, normalization, and multiple co-linearity tests), the LM tests for both spatial lags and spatial errors rejected the original hypothesis at 1% significance level, confirming the necessity of introducing geospatial elements into the regression model. The study then evaluated whether SDM can be reduced to SLM or SEM using the LR test. The results showed that both LR-SLM and LR-SEM rejected the null hypothesis of "no spatial lag" and "no spatial error" at 1% confidence level. Based on the preprocessing results, the SDM model was selected as the best fitting model to evaluate the driving effect of urban land expansion on habitat quality. The SDM estimation and test results are shown in Table 5.

Table 5. Empirical results of Ordinary Least Squares (OLS) and spatial Durbin model (SDM).

Variables	Ordinary Least Squares		Spatial Durbin Model	
	2000–2010	2010–2018	2000–2010	2010–2018
ln(*UEI*)	−0.023	−0.123	−0.029	−0.114 **
ln(*PU*)	−0.06	−0.259 **	−0.286 ***	−0.303 ***
ln(*GDP*)	−0.236 ***	−0.006	0.007	0.058
ln(*AFA*)	0.121 *	0.153 ***	0.021	0.073
ln(*SLO*)	0.087	0.131 **	0.257 ***	0.461 ***
ln(*PRE*)	0.066	0.218 ***	0.064	0.265
ln(*NDVI*)	1.223 ***	1.108 ***	−0.389 **	0.239
$W \times$ ln(*UEI*)			−0.01	−0.084
$W \times$ ln(*PU*)			0.008	0.282
$W \times$ ln(*GDP*)			−0.376 ***	−0.300 *
$W \times$ ln(*AFA*)			0.004	0.102
$W \times$ ln(*SLO*)			−0.328 **	−0.529 ***
$W \times$ ln(*PRE*)			−0.021	−0.225
$W \times$ ln(*NDVI*)			1.819 ***	0.364
R^2	0.430	0.681	0.675	0.775
sigma^2	0.014	0.014	0.008	0.008

Note: * statistical significance at 10% level; ** statistical significance at 5% level; *** statistical significance at 1% level.

The regression results show that the effect of urban land expansion on habitat quality is not significant from 2000 to 2010, while the effects of population urbanization level and NDVI are significantly negative and the effect of slope is significantly positive. In addition, compared to OLS regressions, SDM also highlights the negative spillover effects of GDP and slope on habitat quality and the strong positive spillover effects of NDVI under the influence of spatial factors (Table 5, column 3). The impact of urban land expansion on habitat quality was significantly negative from 2010 to 2018, while the negative impact of population urbanization level was rapidly enhanced (Table 5, column 4). While the results lay the foundation and offer important inspiration, they may only provide a basic picture for the "urban land expansion-habitat quality" story. According to LeSage and Pace (2010) and Du et al. (2019), the SDM coefficients neither reflect the marginal nor the total effects of the independent variables. Instead, they require further decomposition into local direct effects, spatial spillover effects, and total effects by partial differential equations (P.D.E) to comprehensively understand the effects of urban land expansion on habitat quality.

Using partial differential equations, the direct (local) and indirect (spillover) effects were decomposed, and the results are shown in Table 6. Except for 2000, urban land expansion has a significant negative impact on habitat quality in both local and adjacent areas, and the intensity of the impact is increasing rapidly. From 2000 to 2010, the local direct effect of urban land expansion is not significant, but the spillover effect on neighboring areas is significantly negative with a regression coefficient of −0.011 (Table 6, columns 1 and 3). The direct and spatial spillover local effects of urban land expansion were significantly negative from 2010 to 2018, with regression coefficients of −0.135 and −0.317, respectively (Table 6, column 2 and 4). There are several important economic implications of these findings. First, the direct effect shifted from insignificant to negative, while the negative impact of spillover effects increased. This suggests that the influence of urban land expansion on habitat degradation in local and adjacent areas has gradually increased, and the reason will be discussed in Section 4. Second, comparing effect intensities, the negative externality of urban land expansion on habitat quality was more prominent in adjacent areas than in local areas. In addition, regardless of the period, the total effect of urban land expansion was negative and significant. This means that the spatial spillover effect dominates over local direct effects, and therefore, as a whole, urban land expansion reduces regional habitat quality (Table 6, columns 5 and 6). In this sense, the results provide evidence for the need to integrate the conservation and restoration of habitats and the regional coordination of urban land expansion in planning and land resource management.

Table 6. Decompositions of the local, spatial spillover, and total effects for the variables.

Variables	Local Direct Effects		Spatial Spillover Effects		Total Effects	
	2000–2010	2010–2018	2000–2010	2010–2018	2000–2010	2010–2018
ln(UEI)	−0.028 (−0.801)	−0.135 ** (−2.011)	−0.011 * (−1.231)	−0.317 * (−1.755)	−0.039 * (−1.326)	−0.452 *** (−3.978)
ln(PU)	−0.286 *** (−4.540)	−0.289 *** (−3.194)	0.004 (0.041)	0.240 (0.605)	−0.281 *** (−2.905)	−0.048 (−0.110)
ln(GDP)	0.005 (−0.628)	0.024 (0.312)	−0.382 *** (−4.118)	−0.548 (−1.404)	−0.377 *** (−4.891)	−0.524 (−1.190)
ln(AFA)	0.020 (0.395)	0.091 * (1.858)	0.003 (0.036)	0.281 (1.527)	0.024 (0.256)	0.373 * (1.825)
ln(SLO)	0.255 ** (2.503)	0.423 *** (4.229)	−0.333 *** (−2.513)	−0.573 *** (−2.942)	−0.077 (−1.103)	−0.149 (−0.914)
ln(PRE)	0.069 (0.457)	0.252 (1.560)	−0.026 (−0.149)	−0.165 (−0.634)	0.042 (0.917)	0.086 (0.540)
ln(NDVI)	−0.379 * (−1.875)	0.306 (1.433)	1.833 *** (8.468)	0.993 *** (3.189)	1.454 *** (20.053)	1.300 *** (5.621)

Note. T values are in parentheses. * statistical significance at 10% level; ** statistical significance at 5% level; *** statistical significance at 1% level.

There were also interesting results regarding the impact of socio-economic factors on habitat quality, particularly the geographically heterogeneous driving effects of population urbanization GDP. Specifically, population urbanization from 2000 to 2018 threaten local habitat quality, but the effect on adjacent areas was not significant. It can be interpreted that the high concentration of urban population increases the demand for local infrastructure and physical resources, thus creating negative effects on natural habitats [36]. GDP did not have a significant effect on local habitat quality in 2000–2010, but significantly weakened habitat quality in adjacent areas. One possible explanation is that although GDP implies natural resource depletion, housing construction, and green space encroachment, environmental regulation policies and green production technologies offset the negative effects on local habitat. However, in the context of regional integration, local economic construction is likely to generate resource siphoning to adjacent areas, threatening the habitat quality of adjacent areas [37]. Moreover, to verify the phenomenon just described, in addition to using the queen contiguity matrix, the SDM based on the other spatial weight matrix were implemented to confirm the validity of the findings. The confirmatory test results suggest that the research findings were robust and that no abrupt change in the above index system was observed, thus ensuring the robustness and reliability of the model. In addition, the R^2 of the queen contiguity estimation results is better, which also shows the suitability of the selection of the matrix.

4. Discussion

This section discusses the potential mechanisms through which urban land expansion affects the habitat quality of the CUA, especially the different levels of impact on the habitat quality of local and adjacent areas. In addition, the research proposes policy applications, shortcomings and future prospects based on the above.

4.1. Spatial Responses of Local and Adjacent Habitat Quality to Urban Land Expansion

Previous studies have argued that urban land expansion has two main effects on habitat quality. First, urban land expansion reduces public green spaces and landscape diversity in urban ecosystems, adversely impacting habitat integrity and environmental self-renewal [38,39]. Second, economic growth and population agglomeration may increase dependence on natural resources and accelerate their depletion. They may increase the growth of impervious areas and vegetation fragmentation, damage urban ecosystem service functions, and decrease the ecological product value, putting greater pressure on ecological habitats [24,40]. In our study, habitat quality was found to have different levels of responses to urban land expansion in local and adjacent areas, highlighting the role of spatial externalities in urban land expansion.

In terms of local direct effects, urban land expansion had a non-significant impact in 2000–2010 and became negative in 2010–2018. The former is consistent with the phenomenon in bivariate spatial autocorrelation analysis that the number of cities with a negative correlation between urban land expansion and habitat quality were relatively few (9.15%) in the early part of the research period. A possible explanation is that given the extensive mountainous forests with excellent ecological bases and high habitat quality in the southwestern, southeastern, and western CUA, the increased difficulties of area development may generate some initial buffer on the adverse environmental effects of urban expansion, resulting in a non-significant regression coefficient [21]. For example, the NDVI in 2018 for the southwestern mountainous cities of Ya'an (slope = 20°) and Leshan (slope = 13°) were 0.84 and 0.83, while the NDVI for central cities of Chengdu (slope = 5°) and Nejiang (slope = 5°) were 0.67 and 0.75. The Lower NDVI and slope are related to urban ecosystems and construction difficulty, which influence the intensity of negative ecological effects of urban land expansion. The latter may be related to the degradation of buffer effect and the increased of urban development intensity. Over time, mountainous areas with high ecological quality may gradually lose their buffer protection from the adverse effects of urban land expansion [41]. Additionally, despite the rapid decline in the UEI indexes of urban agglomerations, demands for urban land expansion and its intensity to habitat transformation would further intensify, supported by modern construction technologies and socio-economic development needs [42,43]. In particular, due to various national strategies, such as China's Western Development Strategy, the Yangtze River Economic Belt, and the construction of the Chengdu-Chongqing Urban Agglomeration, more demand of infrastructure construction and residential housing are developed in the CUA area, increasing the inward utilization intensity of urban lands [21].

In terms of spatial spillover effects, the adverse effects of urban land expansion on habitat quality have significantly increased in adjacent areas. This could be due to the networking of cross-urban linkages and the demonstration effect of urban land development [44]. On the one hand, along with the development of integrated networks of urban agglomerations, supply pressures on building materials (e.g., cement and wood) required for urban construction are released in the surrounding areas due to unified regional markets and differences in natural endowments, threatening the habitat quality of neighboring cities [26,45]. On the other hand, some measures, such as demolition and relocation of rural houses, land transfer, and the removal of counties and establishment of districts by the local governments, may obscure the actual connotation and extent of urban land expansion. These measures promote tax revenue growth and increase urbanization, producing demonstration effects of land finance management within urban agglomerations and contributing to habitat fragmentation and degradation in neighboring cities [46,47]. Given the influence of these two aspects, urban land expansion has a significant adverse impact on the habitat quality of neighboring cities, the intensity of which even exceeds the role of local urban development and construction.

4.2. Policy Implications

Since the first official draft of the Global Biodiversity Framework was issued, countries around the world have made great efforts to maintain domestic species diversity and control habitat quality. The fifteenth meeting of the Conference of the Parties to the Convention on Biological Diversity (COP15) in Kunming brought the attention of international conservation organizations and individuals to China's habitat protection efforts. This study explored the interaction between urban land expansion and habitat quality in the CUA, providing new insights to support optimal regional urban land expansion and habitat conservation management. Based on the research results, we believe that there is still some work to be done for promoting the sustainability of habitat protection in the CUA and the overall optimization of urban land expansion.

First, policymakers and environmental organizations would have to coordinate and balance the needs for urban growth and habitat protection to develop long-term and sustainable development strategies that consider prevailing land use conditions, landscape types, and economic development in the urban agglomerations. Policies and measures should focus on improving the utilization rate of urban land stock, formulating reasonable paths for optimizing urban land expansion based on regional habitat quality levels, promoting more compact spatial layouts and geometric forms of construction lands, and mitigating habitat fragmentation caused by urban land expansion. Urban plans and strategies should be aimed toward the rational and optimal allocation of urban ecological resources and alleviating pressures on environmental resources caused by urban land expansion, economic development, and population agglomeration. For example, strengthening the construction of urban green corridors, open spaces, urban greenways and park cities to achieve unity between the urban expansion and the ecological needs of residents. Second, more attention should be given to the spillover effects of urban land expansion on habitat quality in adjacent areas. Policymakers should strengthen inter-regional capacity for joint action on habitat protection and management based on the "Ecological Environment Protection Plan of Chengdu-Chongqing Twin Cities Economic Circle" and minimize the adverse mediating effects of regional trade on the environment [48–50]. Cross-regional cooperation in the rational allocation and utilization of land resources should be given more attention, especially to strengthen the coordination of regional ecological restoration and management and achieve high-quality ecological city clusters that encourage wellness and are suitable for business.

4.3. Research Limitations and Future Prospects

There are some limitations in this study that should be considered when interpreting the results. First, the threat sources, habitat suitability, and sensitivity of each land-use type in the habitat quality accounting were made uniform and identical and did not consider spatial geographic differences. In the future, the scientificity of the correlation coefficient value set should be improved based on regional ecological surveys and the InVEST operation manual recommendations. Second, this paper focused on the spatio-temporal evolution and prevailing relationships between urban land expansion and habitat quality but did not tackle future development paths. Subsequent studies can implement other approaches (e.g., the CA-Markov model, FLUS and PLUS) to simulate future development scenarios in urban expansion and habitat quality to improve the perspective and practical reference value of the research.

5. Conclusions

In the 21st century, land-use change has become one of the most influential factors in the ecological environment. Urban land expansion is a key link between human activities and the natural environment, and the impact of landscape pattern changes caused by urban land expansion on regional habitat quality cannot be ignored. This paper investigated the spatio-temporal distribution of urban land expansion and habitat quality in the CUA and analyzed the driving effect of urban land expansion on habitat quality from a spatial perspective. The P.D.E method was used to further explore the local direct effects, spatial spillover effects, and total effects of urban land expansion on habitat quality.

The main findings are as follows: (1) Urban lands in the CUA are scattered along the plains and rivers in the middle of the urban agglomeration, increasing by 2890.42 km^2 from 2000 to 2018, mainly due to urban encroachments over farmlands and grasslands. (2) The urban land expansion exhibited a "Core-Periphery" gradient expansion pattern, with core areas situated in the main urban districts of Chengdu and Chongqing. Urban expansion has generally slowed, and the average UEI index dropped to 7.64% in 2010–2018, decreasing 53.95% compared to 2000–2010. (3) The overall habitat quality dropped to 0.905 by 2018, a decrease of 4.03%. The high-value regions of habitat quality were concentrated in the mountainous cities in the southwest, while two major "ring-shaped decline areas" for habitat quality were formed around the main urban districts of Chengdu and Chongqing. (4) "Low-High" (with negative correlation) and "Low-Low" (with positive correlation) clusters were the main associations between urban land expansion and habitat quality changes. In 2010–2018, county units with negative correlation cluster types have significantly increased, indicating that the dichotomy between urban expansion and habitat quality has become more pronounced. (5) The impact of urban land expansion on local habitat quality shifted from insignificant to negative, while the negative externalities on the habitat quality of adjacent areas have been continuously enhanced. Differential effects of population urbanization level and GDP on habitat quality in local and adjacent areas were also found in this investigation.

Author Contributions: Conceptualization, S.W. and G.W.; methodology, S.W. and G.W.; software, S.W. and G.W.; validation, G.W. and F.L.; formal analysis, S.W. and G.W.; investigation, S.W.; resources, G.W.; data curation, G.W.; writing—original draft preparation, S.W. and G.W.; writing—review and editing, F.L.; visualization, S.W. and F.L.; supervision, G.W. and F.L.; project administration, S.W. and G.W.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Guizhou Provincial Science and Technology Projects (Qian Ke He Ji Chu-ZK[2022] General 234).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available in GEE database.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Original meaning of abbreviations.

Abbreviations	**Original Meaning**
UEI	Urban expansion intensity index
PU	Population urbanization level
GDP	Gross Domestic Product
AFA	Agricultural fertilizer application
HQ	Habitat quality
SLO	Slope
PRE	Average annual precipitation
NDVI	Normalized Difference Vegetation Index
SDM	Spatial Durbin Model
SEM	Spatial Error Model
SLM	Spatial Lag Model
LM	Lagrangian Multiplier
LR	Likelihood Ratio estimation
VIF	Variance Inflation Factor
CUA	Chengdu-Chongqing urban agglomeration

References

1. Chase, J.M.; Blowes, S.A.; Knight, T.M.; Gerstner, K.; May, F. Ecosystem decay exacerbates biodiversity loss with habitat loss. *Nature* **2020**, *584*, 238–243. [CrossRef]
2. Nick, M.H.; Lars, A.B.; Jean, C.; Kendi, F.D.; Andrew, G.; Robert, D.H.; Thomas, E.L.; Joseph, O.S.; Mike, P.A.; Cathy, D.C.; et al. Townshend. Habitat fragmentation and its lasting impact on Earth's ecosystems. *Sci. Adv.* **2015**, *1*, e1500052.
3. Newbold, T.; Hudson, L.N.; Hill, S.L.L.; Contu, S.; Lysenko, I.; Senior, R.A.; Börger, L.; Bennett, D.J.; Choimes, A.; Collen, B.; et al. Global effects of land use on local terrestrial biodiversity. *Nature* **2015**, *520*, 45–50. [CrossRef]
4. Jasper, V.V. Direct and indirect loss of natural area from urban expansion. *Nat. Sustain.* **2019**, *2*, 755–763.
5. Gao, J.; O'Neill, B.C. Mapping global urban land for the 21st century with data-driven simulations and Shared Socioeconomic Pathways. *Nat. Commun.* **2020**, *11*, 2302. [CrossRef]

6. Li, G.D.; Fang, C.L.; Li, Y.J.; Wang, Z.B.; Sun, S.; He, S.W.; Qi, W.; Bao, C.; Ma, H.T.; Fan, Y.P.; et al. Global impacts of future urban expansion on terrestrial vertebrate diversity. *Nat. Commun.* **2022**, *13*, 1628. [CrossRef] [PubMed]
7. Sun, L.Q.; Chen, J.; Li, Q.L.; Huang, D. Dramatic uneven urbanization of large cities throughout the world in recent decades. *Nat. Commun.* **2020**, *11*, 5366. [CrossRef]
8. Bai, L.M.; Feng, X.H.; Sun, R.F.; Gao, Z. Spatial and temporal response of habitat quality to urbanization: A case study of Changchun City. *J. Appl. Ecol.* **2020**, *31*, 1267–1277.
9. Yang, Y.Y. Evolution of habitat quality and association with land-use changes in mountainous areas: A case study of the Taihang Mountains in Hebei Province, China. *Ecol. Indic.* **2021**, *129*, 107967. [CrossRef]
10. Wang, B.X.; Cheng, W.M. Effects of Land Use/Cover on Regional Habitat Quality under Different Geomorphic Types Based on InVEST Model. *Remote Sens.* **2022**, *14*, 1279. [CrossRef]
11. Feng, S.; Sun, R.H.; Chen, L.D. Research on the temporal and spatial evolution of habitat quality in Beijing based on changes in land use patterns. *Acta Ecol. Sin.* **2018**, *38*, 4167–4179.
12. Dai, Y.Z.; Li, J.F.; Yang, J.X. Spatial and temporal responses of habitat quality in Changsha metropolitan area to urban expansion. *Adv. Geogr. Sci.* **2018**, *37*, 1340–1351.
13. Li, X.; Liu, Z.S.; Li, S.J.; Li, Y.X. Multi-Scenario Simulation Analysis of Land Use Impacts on Habitat Quality in Tianjin Based on the PLUS Model Coupled with the InVEST Model. *Sustainability* **2022**, *14*, 6923. [CrossRef]
14. Liu, S.S.; Liao, Q.P.; Xiao, M.Z.; Zhao, D.Y.; Huang, C.B. Spatial and Temporal Variations of Habitat Quality and Its Response of Landscape Dynamic in the Three Gorges Reservoir Area, China. *Int. J. Environ. Res. Public Health* **2022**, *19*, 3594. [CrossRef] [PubMed]
15. Gao, L.N.; Tao, F.; Liu, R.R.; Wang, Z.L.; Leng, H.J.; Zhou, T. Multi-scenario simulation and ecological risk analysis of land use based on the PLUS model: A case study of Nanjing. *Sustain. Cities Soc.* **2022**, *85*, 104055. [CrossRef]
16. Hu, Y.G.; Xu, E.; Kim, G.; Liu, C.; Tian, G.H. Response of Spatio-Temporal Differentiation Characteristics of Habitat Quality to Land Surface Temperature in a Fast Urbanized City. *Forests* **2021**, *12*, 1668. [CrossRef]
17. Zhang, X.R.; Song, W.; Lang, Y.Q.; Feng, X.M.; Yuan, Q.Z.; Wang, J.T. Land use changes in the coastal zone of China's Hebei Province and the corresponding impacts on habitat quality. *Land Use Policy* **2020**, *99*, 104957. [CrossRef]
18. Matteo, G.; Andrea, B.; Philippe, B.; Philippe, M.; Jean, O.G.; Pierre, M. Linking fisheries to land use: How anthropogenic inputs from the watershed shape fish habitat quality. *Sci. Total Environ.* **2020**, *717*, 135377.
19. Nathaniel, S.P.; Adeleye, N. Environmental preservation amidst carbon emissions, energy consumption, and urbanization in selected African countries: Implication for sustainability. *J. Clean. Prod.* **2021**, *285*, 125409. [CrossRef]
20. LeSage, J.; Pace, R.K. *Introduction to Spatial Econometrics*; Taylor and Francis: Abingdon, UK; CRC Press: Boca Raton, FL, USA, 2010.
21. Wang, S.C.; Sun, P.J.; Sun, F.; Jiang, S.N.; Zhang, Z.M.; Wei, G.E. The Direct and Spillover Effect of Multi-Dimensional Urbanization on PM2.5 Concentrations: A Case Study from the Chengdu-Chongqing Urban Agglomeration in China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 10609. [CrossRef]
22. He, J.X.; Hu, S. Ecological efficiency and its determining factors in an urban agglomeration in China: The Chengdu-Chongqing urban agglomeration. *Urban Clim.* **2022**, *41*, 101071. [CrossRef]
23. Kang, S.J.; Choi, J.; Yoon, H. Changes in the extent and distribution of urban land cover in the Democratic People's Republic of Korea (North Korea) between 1987 and 2010. *Land Degrad. Dev.* **2019**, *30*, 2009–2017. [CrossRef]
24. Gomes, E.; Inácio, M.; Bogdzevič, K.; Kalinauskas, M.; Karnauskaitė, D.; Pereira, P. Future scenarios impact on land use change and habitat quality in Lithuania. *Environ. Res.* **2021**, *197*, 111101. [CrossRef] [PubMed]
25. Fang, L.; Wang, L.; Chen, W. Identifying the impacts of natural and human factors on ecosystem service in the Yangtze and Yellow River Basins. *J. Clean. Prod.* **2021**, *314*, 127995. [CrossRef]
26. Yohannes, H.; Soromessa, T.; Argaw, M. Spatio-temporal changes in habitat quality and linkage with landscape characteristics in the Beressa watershed, Blue Nile basin of Ethiopian highlands. *J. Environ. Manag.* **2021**, *281*, 111885. [CrossRef] [PubMed]
27. Zhu, C.M.; Zhang, X.L.; Zhou, M.M.; He, S.; Gan, M.Y.; Yang, L.X.; Wang, K. Impacts of urbanization and landscape pattern on habitat quality using OLS and GWR models in Hangzhou, China. *Ecol. Indic.* **2020**, *117*, 106654. [CrossRef]
28. Yuan, H.H.; Wu, D.D.; Ma, R.F.; Jin, Y.; Zhang, Y. Spatial-Correlation between Agglomeration of Cultural &Creative Industries and Urban Built Environment Field in Hangzhou. *Econ. Geogr.* **2018**, *38*, 123–132.
29. Du, Y.Y.; Wan, Q.; Liu, H.M.; Kapsar, K.; Peng, J. How does urbanization influence PM2.5 concentrations? Perspective of spillover effect of multi-dimensional urbanization impact. *J. Clean. Prod.* **2019**, *220*, 974–983. [CrossRef]
30. Lin, H.; Wang, X.Y.; Bao, G.; Xiao, H.J. Heterogeneous Spatial Effects of FDI on CO2 Emissions in China. *Earth's Future* **2022**, *10*, e2021EF002331.
31. Srikanta, S.; Sandeep, B.; Shahid, R.; Saikat, K.P.; Somnath, S. Estimating global ecosystem service values and its response to land surface dynamics during 1995-2015. *J. Environ. Manag.* **2018**, *223*, 115–131.
32. Cui, G.; Zhang, Y.; Shi, F. Study of Spatiotemporal Changes and Driving Factors of Habitat Quality: A Case Study of the Agro-Pastoral Ecotone in Northern Shaanxi, China. *Sustainability* **2022**, *14*, 5141. [CrossRef]
33. Wang, H.; Hu, Y.; Yan, H. Trade-off among grain production, animal husbandry production, and habitat quality based on future scenario simulations in Xilinhot. *Sci. Total Environ.* **2022**, *817*, 153015. [CrossRef] [PubMed]

34. Wei, G.E.; Zhang, Z.K.; Ouyang, X.; Shen, Y.; Jiang, S.N.; Liu, B.L.; He, B.J. Delineating the spatial-temporal variation of air pollution with urbanization in the Belt and Road Initiative area. *Environ. Impact Assess. Rev.* **2021**, *91*, 106646. [CrossRef]
35. Sun, X.; Jiang, Z.; Liu, F. Monitoring spatio-temporal dynamics of habitat quality in Nansihu Lake basin, eastern China, from 1980 to 2015. *Ecological Indicators.* **2019**, *102*, 716–723. [CrossRef]
36. Cheba, K.; Kiba, J.M.; Baraniecka, A.; Kołakowski, T. Impact of external factors on e-commerce market in cities and its implications on environment. *Sustain. Cities Soc.* **2021**, *72*, 103032. [CrossRef]
37. Rukya, S.; Abduani, A.; Li, H.; Nigati, K.; Li, X.H. Monitoring and Evaluation of Dynamic Changes of Ecological Environment in Fukang City Based on Remote Sensing Ecological Index Model. *Soil Water Conserv. Res.* **2020**, *27*, 283–297.
38. Hu, Y.G.; Xu, E.K.; Dong, N.L.; Tian, G.H.; Kim, G.; Song, P.H.; Ge, S.D.; Liu, S.D. Driving Mechanism of Habitat Quality at Different Grid-Scales in a Metropolitan City. *Forests* **2022**, *13*, 248. [CrossRef]
39. Zhang, X.; Lyu, C.J.; Fan, X.; Bi, R.T.; Xia, L.; Xu, C.C.; Sun, B.; Li, T.; Jiang, C.G. Spatiotemporal Variation and Influence Factors of Habitat Quality in Loess Hilly and Gully Area of Yellow River Basin: A Case Study of Liulin County, China. *Land* **2022**, *11*, 127. [CrossRef]
40. Mengist, W.; Soromessa, T.; Feyisa, G.L. Landscape change effects on habitat quality in a forest biosphere reserve: Implications for the conservation of native habitats. *J. Clean. Prod.* **2021**, *329*, 129778. [CrossRef]
41. Suraj, U.; Puneet, D. Conversion of forestlands to blueberries: Assessing implications for habitat quality in Alabaha river watershed in Southeastern Georgia, United States. *Land Use Policy* **2019**, *89*, 104229.
42. Wen, X.; Liu, Z.H.; Lei, X.H.; Lin, R.J.; Fang, G.H.; Tan, Q.F.; Wang, C.; Tian, Y.; Quan, J. Future changes in Yuan River ecohydrology: Individual and cumulative impacts of climates change and cascade hydropower development on runoff and aquatic habitat quality. *Sci. Total Environ.* **2018**, *633*, 1403–1417. [CrossRef]
43. Marta, T.; Sergi, S.; Becky, C.K.; Lisa, M.; Guy, Z.; Vicenç, A. Model development for the assessment of terrestrial and aquatic habitat quality in conservation planning. *Sci. Total Environ.* **2016**, *540*, 63–70.
44. Tang, F.; Fu, M.; Wang, L.; Zhang, P.T. Land-use change in Changli County, China: Predicting its spatio-temporal evolution in habitat quality. *Ecol. Indic.* **2020**, *117*, 106719. [CrossRef]
45. Weber, D.; Schaepman, S.G.; Ecker, K. Predicting habitat quality of protected dry grasslands using Landsat NDVI phenology. *Ecol. Indic.* **2018**, *91*, 447–460. [CrossRef]
46. Li, F.; Wang, L.; Chen, Z.J.; Clarke, K.C.; Li, M.C.; Jiang, P.H. Extending the SLEUTH model to integrate habitat quality into urban growth simulation. *J. Environ. Manag.* **2018**, *217*, 486–498. [CrossRef]
47. Yang, H.; Huang, Q.; Zhang, J.; Songer, M.; Liu, J. Range-wide assessment of the impact of China's nature reserves on giant panda habitat quality. *Sci. Total Environ.* **2021**, *769*, 145081. [CrossRef]
48. Yang, L.; Liang, Y.; He, B.; Lu, Y.; Gou, Z. COVID-19 effects on property markets: The pandemic decreases the implicit price of metro accessibility. *Tunn. Undergr. Space Technol.* **2022**, *125*, 104528. [CrossRef]
49. Yang, L.; Yu, B.; Liang, P.; Tang, X.; Li, J. Crowdsourced data for physical activity-built environment research: Applying Strava data in Chengdu, China. *Front. Public Health* **2022**, *10*, 883177. [CrossRef]
50. Xu, X.; Wang, Q.; Li, C. The Impact of Dependency Burden on Urban Household Health Expenditure and Its Regional Heterogeneity in China: Based on Quantile Regression Method. *Front. Public Health* **2022**, *10*, 876088. [CrossRef]

sustainability

Article

Self-Sufficiency of New Administrative Capitals (NACs) Based on Types and Commuting Characteristics of Citizens: Case Study of Sejong

Jin-hong Im and Seongman Jang *

Department of Urban and Regional Development, Mokpo National University, Muan-gun 58554, Korea
* Correspondence: sjang@mnu.ac.kr

Abstract: In recent decades, new administrative capitals (NACs) are being developed in Asia and developing countries due to the overcrowding of capitals and large cities. The self-sufficiency of a planned city is considered important for balanced national development. However, no study has specifically analyzed the degree of self-sufficiency of NACs. Therefore, focusing on the city of Sejong (NAC, South Korea) as an example, this study evaluated self-sufficiency using data regarding household composition and travel characteristics. The results of the three-step analysis are as follows: First, the commuting distance of the NAC was longer than that of traditionally developed cities, with relatively little internal commuter traffic in the NAC. Second, commuting to and from the NAC was primarily to large cities nearby. Third, regarding the characteristics of households living in the NAC, the ratio of second-generation households was higher and that of single-person households was relatively small compared with traditional cities. In addition, a spatial correlation in the form of a longer commuting distance in the second generation and shorter commuting distance in single-person households was confirmed. The findings of this study hold important implications for policymakers and urban planning bodies when developing an NAC.

Keywords: new administrative capitals; self-sufficiency; travel characteristics; household composition; commuting characteristics

check for updates

Citation: Im, J.-h.; Jang, S. Self-Sufficiency of New Administrative Capitals (NACs) Based on Types and Commuting Characteristics of Citizens: Case Study of Sejong. *Sustainability* **2022**, *14*, 13193. https://doi.org/10.3390/su142013193

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 31 August 2022
Accepted: 6 October 2022
Published: 14 October 2022

1. Research Background and Objectives

New administrative capitals (NACs) are being developed worldwide to allay the side-effects of overpopulation in big cities and generally pursue balanced national development.

Sejong is one of South Korea's NACs. Sejong City was selected as a candidate for the new administrative capital as it received a high evaluation in five categories: balanced national development effect, domestic and foreign accessibility, impact on the natural environment, natural conditions as a living ground, and urban development cost/economic feasibility. On 11 August 2004, Sejong was confirmed as the target area for the new administrative capital after collecting public opinion through public hearings and various media in 13 cities, including Seoul [1].

As part of the Policy on Balanced National Development, Sejong was designed to reduce overpopulation in metropolitan areas. Initially, the objective was to create an administrative capital by relocating government offices and administrative agencies to Sejong. Sejong was, however, labeled an administrative city since it was considered unconstitutional to build a new administrative capital [2]. Although Sejong is referred to as an administrative city instead of an administrative capital, it functions as a de facto administrative capital, and most administrative agencies have relocated there.

The self-sufficiency of cities is a widely known and acknowledged concept, with several studies conducted on the topic [3,4]. Self-sufficiency is an essential function of cities in terms of their maintenance and construction. The self-sufficiency of cities is critical for an NAC designed for balanced national development [5].

Although several studies have been conducted on policy-level and social aspects of NACs [6–8], there exists insufficient research evaluating whether NACs are self-sufficient. Accordingly, this study aimed to analyze the self-sufficiency of Sejong as an NAC.

2. Literature Review

2.1. Literature on NACs

In pursuit of balanced national development, many NACs have recently been built worldwide by relocating various government departments and administrative agencies and investing massive resources. An administrative capital refers to the central city responsible for the main administrative functions of a country.

South Korea has built Sejong, an NAC, to allay the overcrowding of Seoul, the capital city, and pursue the national policy of balanced national development [9].

Egypt has built an NAC approximately 45 km east of Cairo to resolve issues emerging from overpopulation in its capital. Key government organs, including parliament, have been relocated to the administrative capital. Several studies have centered on this development [7,10,11].

In the Philippines, a state-led urban development plan to build an administrative capital (New Clark City) has emerged to decentralize the population in the metropolitan areas. A study has focused on sustainable environmental city research on the topic [12].

In Malaysia, Putrajaya was built as the administrative capital as a remedy to combat overpopulation and traffic congestion in its capital, Kuala Lumpur [13]. Various studies on urban growth and ecosystem protection in Putrajaya have been conducted [14].

Indonesia is pursuing a state-led urban planning project to relocate the capital from Jakarta to the eastern part of Kalimantan. The Indonesian parliament passed the Capital City Bill on 18 January 2022, leading to many studies on urban planning in relation to capital relocation [8,15].

Therefore, NAC construction is currently a global trend. Countries that have completed or are building and discussing NACs are shown in Table 1. In addition to building new administrative capitals, many countries have pursued state-led planned city development, such as capital relocation and large-scale new town development for various reasons, including addressing environmental problems and efficiently utilizing urban space. Such forms of urban development are commonly found in Asia and developing nations.

The Xiong'an New Area near Beijing in China is a state-led project designed to disseminate key functions of the Chinese capital. As one of President Xi Jinping's core projects, the project was designed and developed as a national-level special zone, receiving an investment of over USD 300 billion. Various related urban planning studies have been conducted [16,17].

State-led planned city construction is being promoted in many countries in addition to China, and the construction of new capitals has been completed in various countries, including Pakistan (Islamabad) and Kazakhstan (Astana) [9].

Table 1. Cases of construction and discussion of NAC construction.

Classification	Brazil	Australia	Japan	Malaysia	Republic of Korea	Egypt	Philippines	Indonesia
Type	NAC Construction	NAC Construction	NAC Construction	NAC Construction	NAC Construction	NAC Construction	NAC Discussion	NAC Discussion
Purpose/ Background	Inland area development	Federal national symbol project	Balanced national development	Balanced national development	Balanced national development	Balanced national development	Balanced national development	Balanced national development, resolving environmental issues
Period	1955–1970	1908–1980	1992–Currently invalid	1993–2010	2007–Currently in progress	2015–Currently in progress	Currently under discussion	Currently under discussion
Special note	Security and national development center	Academic/research/art function integration	Reflecting the national characteristics of decentralization	Except for the Royal Family and the Bundestag	Changed to a multi-functional administrative city due to unconstitutional issues	5 million metropolis targets	Considered as a way to spread the growth of large cities	President's strong push

2.2. Literature on Self-Sufficiency

Several studies have been conducted on the self-sufficiency of cities, especially new cities [3,4]. Self-sufficiency is an important factor in the functionality of cities. In particular, the self-sufficiency of cities has become a key factor for NACs, which are created with the goal of balanced national development and the relocation of capital cities.

Urban self-sufficiency is an important function that allows new cities to become more sustainable [18]. A study identified self-sufficiency as a key criterion for sustainability indices [19]. Some studies have focused on the "land" aspect in evaluating self-sufficiency. One study focused on establishing self-sufficiency through housing site/lot development projects [20], and another stressed establishing self-sufficiency through the allocation of land to enhance the self-functionality of a city [21].

Commutes are used as a primary indicator to evaluate urban self-sufficiency. A study revealed that commuting is an important indicator when evaluating the economic self-sufficiency of a city [22]. Another study conducted a comparative analysis of the level of economic self-sufficiency in new cities by focusing on commutes [23]. While evaluating the self-sufficiency of new cities (Phase 1) in metropolitan areas, commutes were confirmed as a key indicator [24]. A study derived self-sufficiency variables and used the analytic hierarchy process (AHP) to conduct a time-series analysis of self-sufficiency in metropolitan areas [25]. The authors concluded that economic self-sufficiency can be determined through the efficiency of commutes. Another study argued that commuting to work reflects the economic function of a city and that the clarity of origin-destination location and regular travel can be sufficient indicators for measuring the evaluation index of economic self-sufficiency [26].

Reducing commuting distances is a prerequisite for cities to become self-sufficient and furthermore to build sustainable cities. For urban self-sufficiency, a work—worker balance must be sought. Some studies found that better work—worker balance resulted in shorter average commuting distances and times [27–29]. A study claimed that shorter commutes, both in terms of time taken and distance, are desirable for greater sustainability [30]. Another study emphasized that commuting distance can be reduced depending on the land use type and that reducing the commuting distance will lead to the expansion of a sustainable city [31].

The results of the above studies have confirmed that urban self-sufficiency is critical for the optimal development of cities. In addition, commutes have been identified as a key indicator to evaluate the self-sufficiency of cities. Accordingly, this study used commutes as a key indicator to evaluate the self-sufficiency of Sejong.

While several studies have been conducted on the policies and development of new cities [32,33], there is insufficient research on the self-sufficiency of cities, not-withstanding the fact that NACs have been built in line with comprehensive balanced national development plans. Therefore, this study assessed the self-sufficiency of Sejong using indicators used in previous studies. The study is unique in that it evaluated self-sufficiency using correlation analysis between travel behavior and household composition.

3. Research Scope and Methods

3.1. Research Scope

The administrative districts of South Korea are composed of one special city, six metropolitan cities, eight provinces, one special self-governing province, and one special self-governing city. The subject of this study is Sejong, an NAC in South Korea (Figure 1). Sejong was initially designed as an administrative capital following the relocation of government offices and administrative agencies [2]. It was, however, labeled an administrative city based on constitutional requirements, as stated earlier.

The full-scale development of Sejong began in 2007 with the goal of building a self-sufficient city with a population of 500,000 by 2030 aimed at pursuing balanced national development and enhancing national competitiveness [34]. The development of Sejong has been conducted in three phases. Phase 1 (initial phase), from 2007 to 2015, involved the relocation of the central administrative agency and the building of urban infrastructure. Phase 2 (maturity phase), from 2016 to 2020, involved fostering self-sufficiency and developing urban infrastructure. Phase 3 (completion phase), from 2021 to 2030, involves the completion of a self-sufficient city.

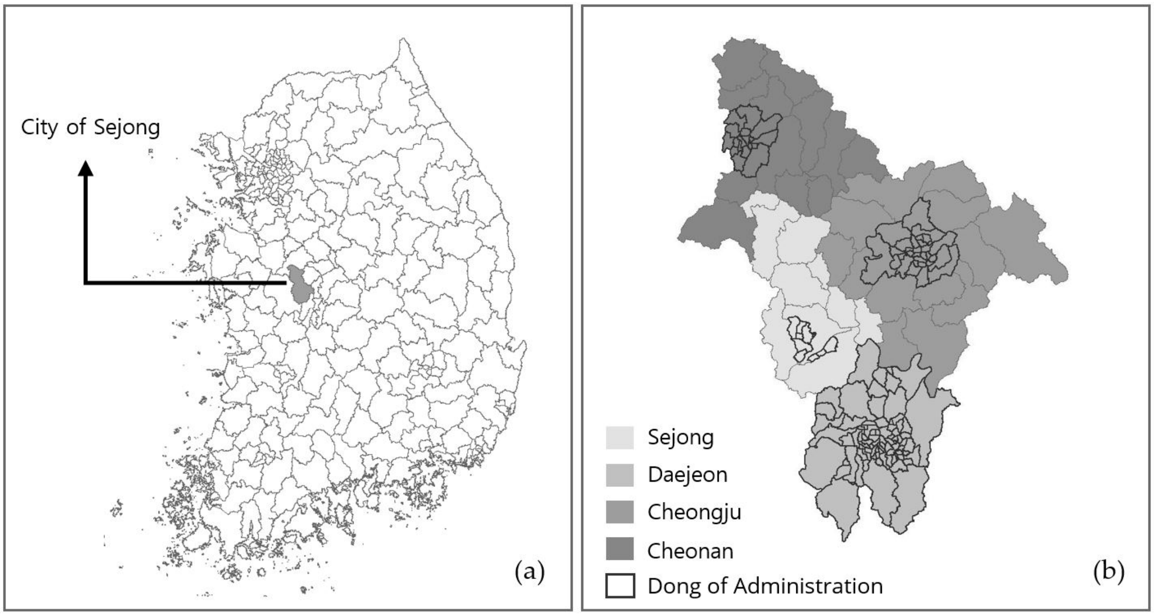


Figure 1. Sejong, South Korea's NAC (**a**), in comparison to other cities (**b**), Daejeon, Cheongju, and Cheonan.

This study evaluated the self-sufficiency of Sejong, which has recently completed its Phase 2 development. The cities analyzed in comparison with Sejong were selected among those geographically close and demographically comparable to Sejong (Daejeon, Cheongju, and Cheonan) (Figure 1). The regional scope of research was narrowed to administrative *dongs*, or neighborhoods in which the new administrative capital is built. Accordingly, the administrative dong was also set as the research scope for the cities used in the comparison.

3.2. Research Method

Similar to previous studies, this study set commutes used in self-sufficiency indices as the key indicator of the self-sufficiency of Sejong. Regarding the data sources, to analyze commutes, we used the Korea Transport Database for data on commutes from 2019, when the Phase 2 development of Sejong was near completion. For geographic information system (GIS)-related analysis, we used the Statistical Regional Boundary (2019) data of the Statistical Geographic Information Service (SGIS). For analysis by household type, we used statistical data from SGIS (2020).

The specific data analysis process was divided into three stages. First, to compare and analyze basic economic self-sufficiency, we extracted the travel distance for each city as well as the ratio of intra-city and inter-city travel and confirmed the same using analysis of variance (ANOVA). Second, we verified the spatial dispersion of the destination of commutes in each region using GIS to analyze the commuting behavior of travelers by city. Third, we verified the correlation between household type and self-sufficiency in cities using bivariate Moran's I, the LISA cluster map, ANOVA, and a spatial economic model.

4. Data Analysis and Results

4.1. Commuting Distance and Ratio of Intra-City and Inter-City Travel

To analyze the commuting distance by city, we obtained commuting data and the SGIS of Sejong and the cities for comparison (Daejeon, Cheongju, and Cheonan). We then calculated the distance between dongs using GIS. When the point of departure and destination are the same, the distance is considered zero. In this case, the dong area with the same origin and destination is set as S. The area (S) of this region was assumed to be a circle, and the r value was set as the distance value. The value of r was obtained using $r = \sqrt{\frac{S}{\pi}}$. The following Equation (1) was used to obtain the average commuting distance by city. The results are shown in Figure 2.

$$AD_C^P = \frac{\sum_i^n \frac{T_{c_{ij}}^p \times D_{ij}}{\sum_i^n T_{c_i}^p}}{n} \quad (1)$$

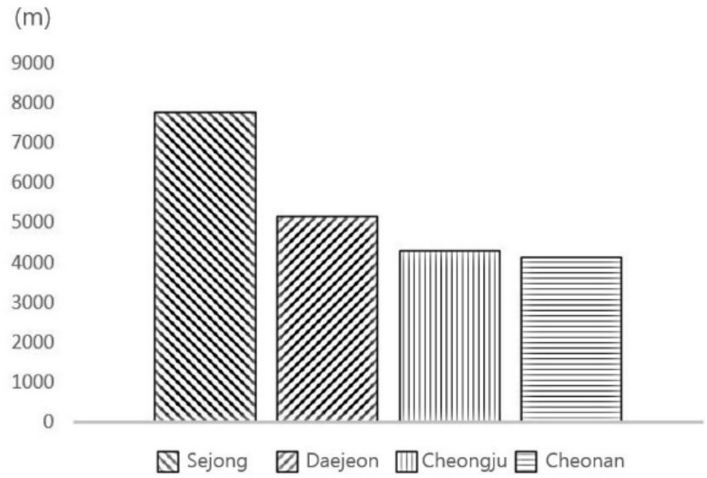


Figure 2. Average commuting distance by city.

AD_C^P is the average travel distance of P in city C; $T_{c_{ij}}^p$ is the total time for P to travel from the starting point i to the ending point j of city C; D_{ij} is the distance from the starting point i to j; C is the city analyzed (Sejong, Daejeon, Cheonan, or Cheongju); P is the purpose of travel (commute to work, commute to school, commute to a private academy, or commute to shopping areas); n is the number of administrative dongs in city C.

We identified the commutes in dongs of each city to analyze intra-city and inter-city commutes. The destinations of the commutes were identified from all areas, including the dongs of Sejong and the other cities compared. We calculated the ratio of intra-city travel by processing the data examined.

We conducted ANOVA to assess whether the differences in the average travel distance among the cities and the ratios of intra-city travel were statistically significant. In addition, we conducted multiple comparisons using the Tukey test. The results are shown in Tables 2 and 3.

The data analysis in Table 2 shows that the difference in the average commuting distance between Sejong and the compared cities was 2621 for Daejeon, 3472 for Cheongju, and 3634 for Cheonan, with a p value of $p < 0.001$. Therefore, the average distance of commutes in Sejong was statistically significantly greater than that of the compared cities.

The results of the analysis in Table 3 show that the ratio of intra-city commutes and the ratio of trips made for shopping within the city had negative values ($-$) compared to those of the other cities. The mean difference between the ratio of intra-city commutes and the ratio of trips made for shopping within the city showed a p value of $p < 0.001$. This indicates that the ratio of intra-city commutes and the ratio of trips made for shopping within the city of Sejong were lower than those of the other cities. Therefore, Sejong has a higher ratio of inter-city travel for commuting to work and shopping.

The above multiple comparisons indicate that, first, travelers in Sejong, on average, commute for a longer distance than those in other cities. Second, travelers in Sejong engage in a greater ratio of inter-city travel for commuting to work and for shopping than those in other cities. Although self-sufficiency in a new administrative capital is an important urban function, according to the results of the analysis, Sejong does not yet meet the standards of a self-sufficient city.

Table 2. Multiple comparisons of average commutes by city using the Tukey test.

Multiple Comparisons							
Tukey's HSD							
Dependent variable	(i) Regional code	(j) Regional code	Mean difference (i–j)	Standard error	*p* value	95% confidence interval	
						Lower bound	Upper bound
Average commuting distance	Sejong	Daejeon	2621.0023636 *	588.0476634	0.000	1090.866767	4151.137960
		Cheongju	3472.1191656 *	635.2671735	0.000	1819.115547	5125.122784
		Cheonan	3634.0101259 *	682.3864079	0.000	1858.399409	5409.620843
Average distance to school	Sejong	Daejeon	−983.7157282	852.7424699	0.657	−3202.603225	1235.171768
		Cheongju	−2415.3300036 *	921.2166501	0.048	−4812.391460	−18.268548
		Cheonan	−1373.0647563	989.5454180	0.509	−3947.921800	1201.792288
Average distance to private academy	Sejong	Daejeon	−624.1560903	486.3043008	0.575	−1889.549302	641.237121
		Cheongju	−594.5800376	525.3539429	0.671	−1961.582780	772.422705
		Cheonan	−1255.5413656	564.3206590	0.122	−2723.937861	212.855130
Average distance to shopping areas	Sejong	Daejeon	514.5920831	494.6446715	0.726	−772.503278	1801.687444
		Cheongju	637.8508508	534.3640352	0.632	−752.596697	2028.298399
		Cheonan	820.1581663	573.9990507	0.484	−673.422090	2313.738422

* The mean difference is significant at the level of 0.05.

Table 3. Multiple comparisons of average internal ratio by city using the Tukey test.

Multiple Comparisons							
Tukey's HSD							
Dependent variable	(I) CODE	(J) CODE	Mean difference (I–J)	Standard error	*p* value	95% confidence interval	
						Lower bound	Upper bound
Internal commuting ratio	Sejong	Daejeon	−0.1948380 *	0.0208854	0.000	−0.249183	−0.140493
		Cheongju	−0.2060066 *	0.0225625	0.000	−0.264716	−0.147298
		Cheonan	−0.2036130 *	0.0242360	0.000	−0.266677	−0.140549
Internal school commuting ratio	Sejong	Daejeon	−0.0341300	0.0189290	0.277	−0.083384	0.015125
		Cheongju	−0.0057187	0.0204490	0.992	−0.058928	0.047491
		Cheonan	−0.0324174	0.0219658	0.455	−0.089574	0.024739
Internal private academy commuting ratio	Sejong	Daejeon	−0.0008355	0.0297607	1.000	−0.078275	0.076604
		Cheongju	−0.0005510	0.0321504	1.000	−0.084208	0.083106
		Cheonan	0.0546093	0.0345351	0.393	−0.035253	0.144472
Internal shopping commuting ratio	Sejong	Daejeon	−0.0545728 *	0.0067486	0.000	−0.072133	−0.037013
		Cheongju	−0.0561269 *	0.0072905	0.000	−0.075097	−0.037157
		Cheonan	−0.0569028 *	0.0078312	0.000	−0.077280	−0.036525

* The mean difference is significant at the level of 0.05.

4.2. Spatial Variance Analysis of Commuting Destinations by Region Using GIS

To illustrate the commutes of Sejong and the other cities with spatial dispersion and analyze commuting behavior, we used the standard deviational ellipse (SDE) function of GIS. This analysis method has been used in various studies to analyze spatial distribution [35,36].

First, we set the dong of each city as the X point and the destination as the Y point and calculated the volume of commutes based on X and Y. Then, we created an ellipsoid using the deviation representing the distribution of this value. The shape of the SDE is shown in Figure 3.

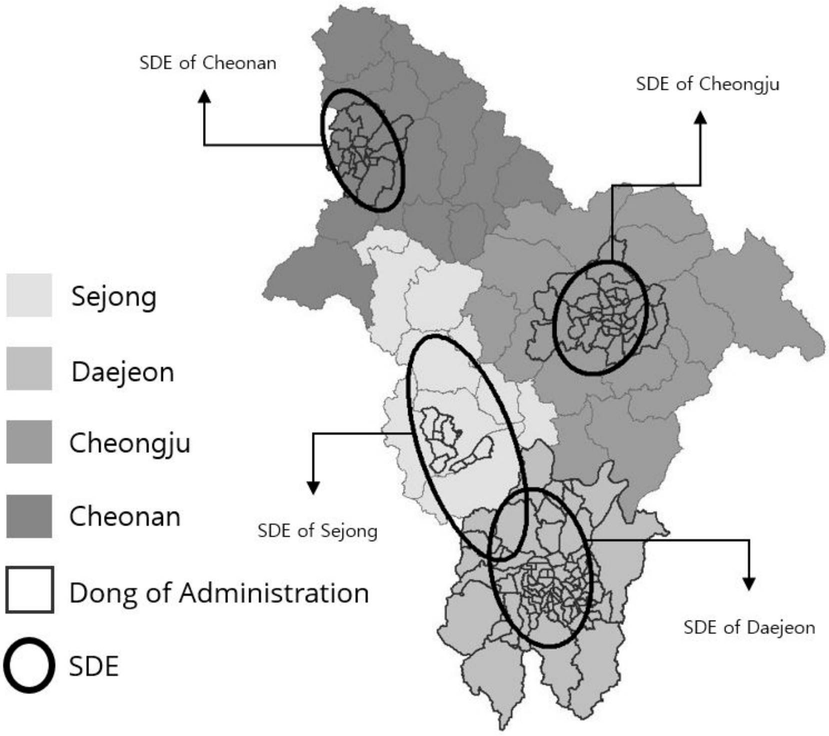


Figure 3. Commuting standard deviational ellipse (SDE) by city.

According to the analysis, the SDE of the compared cities is formed within the area and the ellipsoid is relatively circular. However, the SDE of Sejong is illustrated as an ellipsoid that is wider than that of the other cities. In addition, the ellipsoid moves into that of Daejeon, intersecting it. These results indicate that numerous travelers in Sejong commute to and from the neighboring city Daejeon.

4.3. Analysis of Household Type by Region

Households in South Korea are classified into four types: single-person, one-generational, two-generational, and three-generational households. This classification has been commonly used in previous studies when analyzing household composition [37,38]. We first examined the household member data of Sejong and the other cities to derive the ratio of each household type. The results are shown in Figure 4. Sejong indicated a relatively higher ratio of two-generational households and a lower ratio of single-person households compared to the other cities.

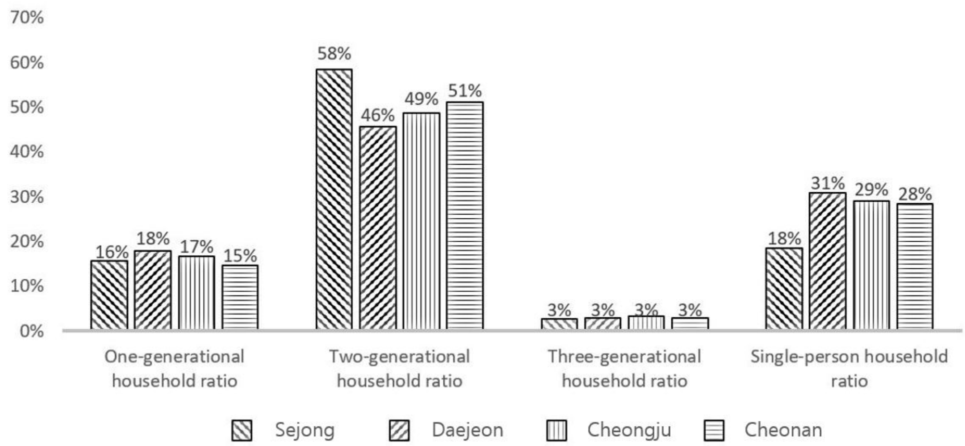


Figure 4. Ratio of household type in the city by *dong*.

To ensure that the values represent meaningful results, we assessed whether the ratios of two-generational households and single-person households in Sejong were statistically significant compared to those of the other cities. We conducted ANOVA on the ratio of household data by city, and multiple comparisons and verification using Tukey's test. The results are shown in Table 4.

Table 4. Multiple comparisons of ratio of household type by city using Tukey's test.

Multiple Comparisons							
Tukey HSD							
Dependent variable	(I) Local_code	(J) Local_code	Mean difference (I–J)	Standard error	p value	95% confidence interval	
						Lower bound	Upper bound
One-generational household ratio	Sejong	Daejeon	−0.0237611 *	0.0030318	0.000	−0.031552	−0.015970
		Cheongju	−0.0113181 *	0.0032984	0.003	−0.019794	−0.002842
		Cheonan	0.0096786 *	0.0034025	0.023	0.000935	0.018422
Two-generational household ratio	Sejong	Daejeon	0.1267152 *	0.0099961	0.000	0.101028	0.152403
		Cheongju	0.0959205 *	0.0108751	0.000	0.067974	0.123867
		Cheonan	0.0727344 *	0.0112183	0.000	0.043906	0.101563
Three-generational household ratio	Sejong	Daejeon	−0.0025234	0.0012602	0.187	−0.005762	0.000715
		Cheongju	−0.0069762 *	0.0013711	0.000	−0.010499	−0.003453
		Cheonan	−0.0021211	0.0014143	0.438	−0.005756	0.001513
Single-person household ratio	Sejong	Daejeon	−0.1246426 *	0.0107597	0.000	−0.152292	−0.096993
		Cheongju	−0.1049691 *	0.0117058	0.000	−0.135050	−0.074888
		Cheonan	−0.0990003 *	0.0120752	0.000	−0.130031	−0.067970

* The mean difference is significant at the level of 0.05.

The results indicated that the mean difference in the ratio of two-generational households between Sejong and the other cities was positive (+), whereas the mean difference in single-person households was negative (−). The mean difference of the ratios of both two-generational and single-person households had a p value of $p < 0.001$. This finding indicates that Sejong comprised more two-generational households and fewer single-person households than the other cities.

To analyze whether the calculated ratio of household composition by city was related to the self-sufficiency of the cities, we evaluated the spatial autocorrelation of the commuting distance and household type ratio by city. Spatial autocorrelation measures the extent to which a variable at a specific location is related to other values at nearby locations. Spatial autocorrelation is positive when the level of interaction exceeds the expected level and the surrounding locations have comparable values. In contrast, it is negative when the high value of one variable is close to the low value of the variable at a nearby location. Spatial autocorrelation is 0 when there is no relationship between close values [39].

We applied bivariate Moran's I and the LISA cluster map, which are the most common techniques to analyze the existence of spatial autocorrelation. Bivariate Moran's I and the LISA cluster map explain the spatial patterns formed by two different variables [40]. Bivariate Moran's I derives the spatial scatter plot of the first variable on the vertical axis and that of the second variable on the horizontal axis. The two variables are internally standardized. Spatial delayed operation is applied to the standardized variables. The slope of the regression line represents the linear correlation between the variable on the horizontal axis and the variable on the vertical axis of a nearby location [41]. Based on Moran's I, the bivariate LISA cluster map provides a feasible method to characterize spatial correlations between the spatial distributions of several variables [42]. Bivariate Moran's I (Ikl) can be presented as shown in Equation (2) below:

$$I_{kl} = Z_k^i \sum_{j=1}^{n} W_{ij} Z_l^j \quad (2)$$

where $Z_k^i = [x_k^i - \overline{x_k}]/\sigma_k$, $Z_l^j = [x_l^i - \overline{x_l}]/\sigma_l$; x_k^i is the value of variable k at location i; x_l^j is the value of variable l at location j; $\overline{x_k}$ and $\overline{x_l}$ are the mean values of the variables k and l, respectively; σ_k and σ_l are the variance of x for variables k and l, respectively; and W_{ij} is the spatial weight matrix, which can be represented based on the distance weighting between locations i and j [39].

First, a bivariate Moran's I scatter plot was derived from the results of the spatial correlation between commuting distance by city and the ratio of household type. The values are shown in Figure 5. The results of the bivariate Moran's I analysis indicate that the correlation between commuting distance by city and the ratio of two-generational and single-person households was higher. The bivariate Moran's I of two-generational households was approximately 0.203, indicating that the regression line was an upward slope. This finding implies that a higher ratio of two-generational households makes it more likely that the commuting distance will be longer. The bivariate Moran's I of single-person households was approximately −0.185, indicating that the regression line was a downward slope. This result implies that a higher ratio of single-person households makes it more likely that the commuting distance will be shorter.

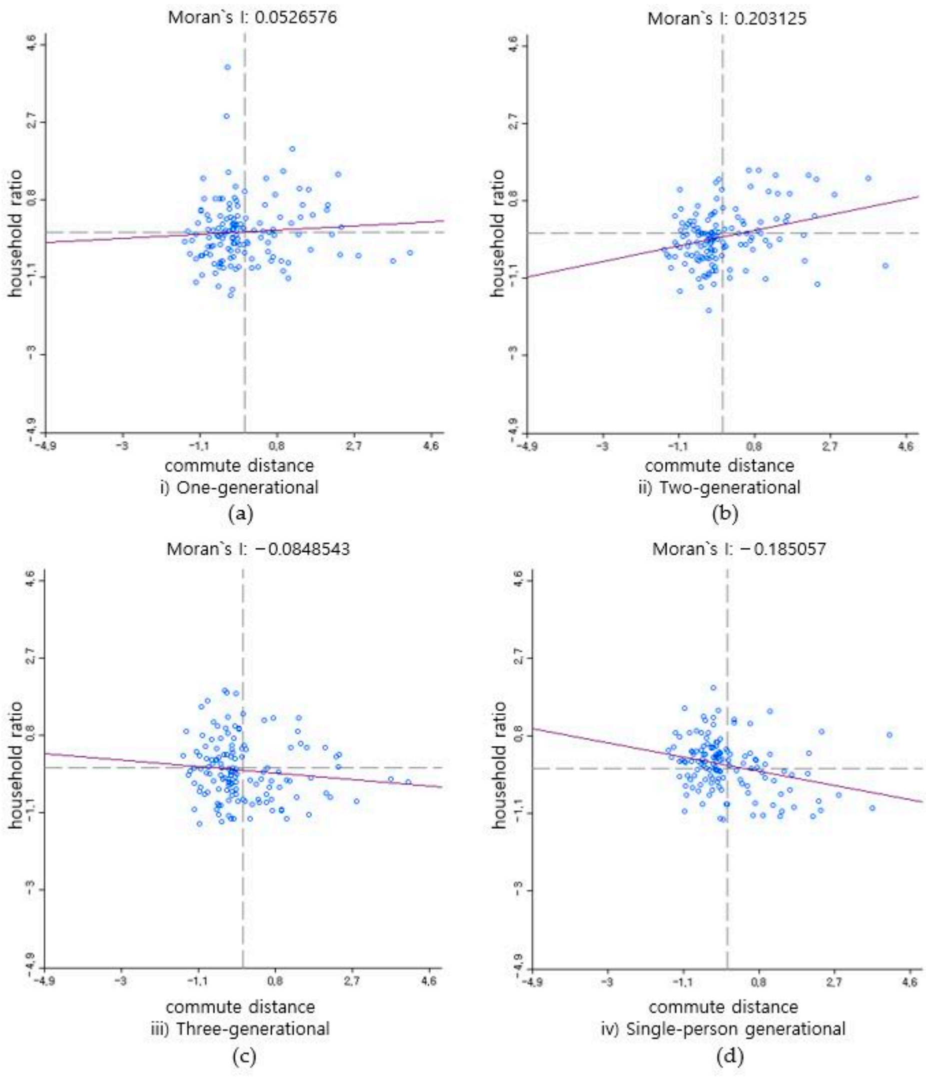


Figure 5. Moran's I and scatter plot of commuting distance and ratio of household type. (**a**) Moran's I and scatter plot of commuting distance and One-generational household. (**b**) Moran's I and scatter plot of commuting distance and Two-generational household. (**c**) Moran's I and scatter plot of commuting distance and Three-generational household. (**d**) Moran's I and scatter plot of commuting distance and Single-person household.

Second, the bivariate LISA cluster map, which illustrates the interaction (spatial correlation) between commuting distance by city and ratio of household type, was classified into six categories: high-high (HH), high-low (HL), low-high (LH), low-low (LL), insignificant variables, and neighborless. The bivariate LISA cluster map between commuting distance by city and the ratio of household type is shown in Figure 6. In particular, in the new administrative capital Sejong, five HH-type regions were derived from the cluster map. The HH-type regions were found only in Sejong. Considering that there are a total of nine dongs in Sejong, approximately 56% of the regions in Sejong were determined to be of the HH type. This finding indicates that Sejong had more two-generational households than the other cities, which suggests a longer average commuting distance, thereby demonstrating a spatial correlation. In addition, in the cluster map of the ratio of single-person households and commuting distance, Sejong indicated a pattern different from the other cities. Four HL-type regions were found only in Sejong. Considering that there are a total of nine dongs in Sejong, approximately 44% of the regions in Sejong were determined to be of the HL type. This finding indicates that fewer single-person households lead to a greater average commuting distance, thereby confirming a spatial correlation.

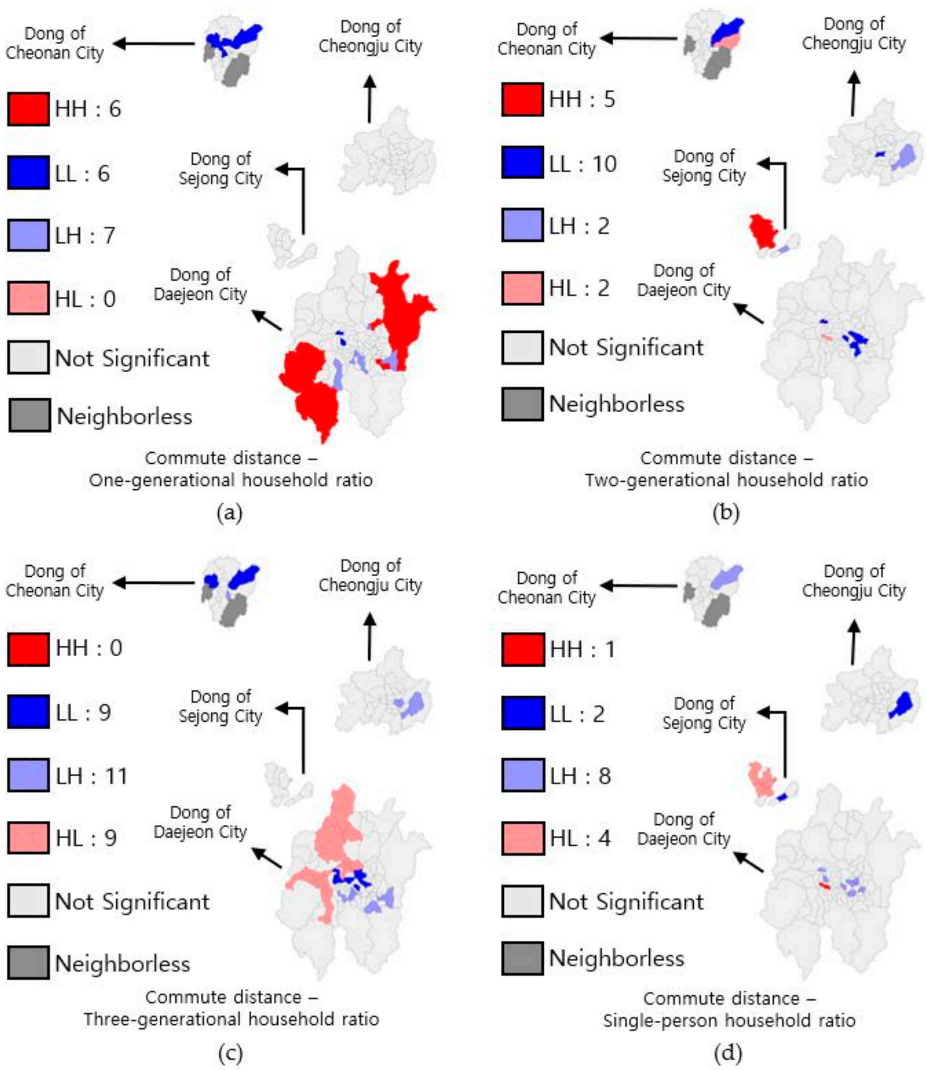


Figure 6. LISA cluster map of household type by city and commuting distance. (**a**) LISA cluster map of commuting distance and One-generational household. (**b**) LISA cluster map of commuting distance and Two-generational household. (**c**) LISA cluster map of commuting distance and Three-generational household. (**d**) LISA cluster map of commuting distance and single-person household.

Therefore, the results of the above analysis indicate that Sejong has more two-generational and fewer single-person households than the other cities. Household type was correlated with commuting distance. The presence of more two-generational households suggests a longer commuting distance on average. In addition, fewer single-person households also indicated a longer commuting distance on average. The new administrative capital was found to comprise more two-generational households than the other cities, which suggested a correlation with longer average commuting distance. Furthermore, the new administrative capital had fewer single-person households than the other cities, also indicating a correlation with longer average commuting distance.

We further analyzed the relationship between household type and commuting distance. An appropriate analysis model was selected to analyze the relationship between the two variables with special autocorrelation. Earlier studies have confirmed that the spatial economic model is suitable to analyze the relationship between variables while controlling autocorrelation [43].

The spatial economic model consists of the following steps. First, the researcher conducts an ordinary least squares (OLS) regression model. If the Lagrange multiplier (LM) statistics of both models as calculated by the LM diagnoses are not significant, OLS is the most appropriate model for the analysis [39]. Secondly, if any of the LM diagnoses from the spatial lag model and spatial error model are meaningful, the corresponding model is determined as an appropriate model. Thirdly, if the LM diagnoses for both models are significant, the robust LM diagnosis is examined, and the model with the relatively higher significance is chosen as the final model.

The appropriate analysis model was chosen using OLS in this study. In OLS, the commuting distance was set as the dependent variable. The independent variable was defined as one-generational, three-generational, single-person, and non-family households, excluding the ratio of two-generational households. Two-generational households were defined as the reference value of the model. In addition, a spatial weight matrix was constructed for LM diagnoses. The standard for the inter-individual adjacency of the spatial weight matrix was applied with the queen-based contiguity standard. Queen-based contiguity is a method of defining the adjacency of a specific spatial entity with all neighboring spatial entities with common edges or corners. Based on the set values, OLS was performed and LM diagnoses were identified. The LM lag value was significant at 0.02083, whereas the LM error value was insignificant at 0.11992. The results of the analysis are shown in Table 5.

Table 5. Results of the spatial dependence diagnoses.

Classification			*p*-Value
Commuting distance	Lagrange multiplier	Lag **	0.02083
		Error	0.11992
	Robust LM	Lag **	0.03289
		Error	0.20188

** The mean difference is significant at the level of 0.05.

Based on these findings, the spatial lag model was selected as the appropriate model for analysis. The derived R-squared value was 0.141762, indicating a somewhat low explanatory power. The results of the analysis are shown in Table 6. The results indicate that the commuting distance decreases when the ratio of single-person households exceeds that of two-generational households. This finding was significant in the standard of significance probability of 5% and is consistent with the analysis results performed previously in this study. Regarding other household types, the commuting distance increased or decreased as the ratio increased. However, because the derived results were insignificant, we did not interpret them further.

Table 6. Analysis of spatial lag model based on commuting distance.

Dependent Variable	Independent Variable	Coefficient	*p*-Value
Commuting distance	Spatial lag Coef **	0.1871784	0.02763
	One-generational household ratio	5007.54	0.41118
	Three-generational household ratio	−37328.51	0.08294
	Single-person household ratio **	−3786.229	0.04884
	Non-family household ratio	−12986.07	0.49267
	R-squared		0.141762

** The mean difference is significant at the level of 0.05.

5. Conclusions

The construction of NACs has emerged as an important developmental goal worldwide [9]. NACs are a new form of urban development without precedents and, therefore, are characteristically different from existing cities. It is important that such new cities are economically self-sufficient to realize balanced development [23]. However, there are a lack of studies that analyze the self-sufficiency of NACs. Regarding previous research, various studies have been conducted on urban self-sufficiency by using commuting as an important indicator. Therefore, this study evaluated the urban self-sufficiency of Sejong, an NAC, using commuting as the main indicator. The analysis indicated the following findings.

First, Sejong lacks self-sufficiency compared to existing cities in terms of commuting distance, which was longer in Sejong. In addition, a higher ratio of inter-city commutes compared to that of intra-city commutes was observed in Sejong. This finding shows that many residents of Sejong commute to other cities, suggesting that the city is not self-sufficient.

Second, Sejong appears to be greatly influenced by neighboring metropolitan cities. The SDE drawn using GIS displayed the travel behavior of people living in Sejong commuting frequently to big cities nearby. It seems that despite the better residential environment of Sejong, its residents commute to neighboring big cities since there are not enough jobs (excluding those at government departments and research institutes that were involuntarily relocated) in Sejong.

Third, Sejong has a different household composition ratio compared with that of other cities. It is presumed that household composition is related to the self-sufficiency of cities. The results of Moran's I analysis showed a spatial correlation in which a higher proportion of two-generational households is linked to greater commuting distance, and fewer single-person households are linked to a longer commuting distance. The results of the LISA cluster map analysis of Sejong also indicated a spatial correlation between two-generational households and long commuting distance, and a spatial correlation between fewer single-person households and long commuting distance. The spatial economic model indicated that the commuting distance decreased as the ratio of single-person households exceeded that of second-generational households. A study found that single-person households are more heavily affected by their work than multi-person households [44]. Based on this finding, it can be assumed that NACs with a low ratio of single-person households that commute to work within a short distance lack a variety of jobs. NACs are unique in that administrative departments and public offices are mostly involuntarily relocated. This may explain the high ratio of two-generational households in the NAC. The fact that household type is related to commuting distance indicates that commuting distance is also related to self-sufficiency.

Based on the above results, the policy implications of NACs are as follows. First, Sejong is not as self-sufficient as it was designed and expected to be. To build a more economically self-sufficient city, policies must attract and relocate firms to the city when the state decides to involuntarily relocate its government departments and research institutes. By gradually and effectively providing government-level benefits such as tax exemptions and subsidies to firms relocating to the NAC, it would be possible to create new jobs within the city. This will encourage single-person households to reduce their commuting distance and help the city become self-sufficient.

Second, the NAC is heavily impacted by the big cities nearby. This raises the need to consider turning the NAC into a megacity and expanding the characteristics of the city. Expanding the metropolitan scale of urban functions to that of a megacity would lead to the development of various industries, the creation of better infrastructure, and the improvement of the residential environment. In this way, the NAC will grow into a city that better meets the objectives of balanced national development. Along with various ongoing studies on megacities [5], the current findings are expected to be effectively used in establishing related policies on NACs.

This study is significant in that it analyzed and verified the self-sufficiency and travel behavior of an NAC and deduced pertinent policy implications. In particular, although many studies have evaluated self-sufficiency using commuting data, there is insufficient research on the correlation between household type and the self-sufficiency of an NAC. Accordingly, this study is novel in that it analyzed the types of household composition, identified their spatial correlation with commuting behavior, and established a correlation between household composition and the self-sufficiency of the NAC.

Although the present study reveals important findings, it has several limitations. The findings of this study are limited in that the self-sufficiency analysis was restricted to Sejong. Further research to evaluate the self-sufficiency of NACs in other countries needs to be conducted to determine whether the results are consistent with those of this study. Moreover, the correlation analysis between household composition and self-sufficiency was also limited to Sejong. Since household composition can be affected by various factors such as gender and age, it is necessary to not only analyze the NACs of other countries but also classify household types using more detailed criteria. Such limitations are hoped to be addressed in future research on the analysis of the correlation between household composition and the commuting characteristics of NACs.

Author Contributions: Conceptualization, S.J.; methodology, S.J.; software, S.J. and J.-h.I.; validation, S.J. and J.-h.I.; formal analysis, S.J.; investigation, J.-h.I.; resources, S.J.; data curation, S.J. and J.-h.I.; writing—original draft preparation, J.-h.I.; writing—review and editing, S.J.; visualization, S.J. and J.-h.I.; supervision, S.J.; project administration, S.J.; funding acquisition, S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2020R1I1A1A01067265).

Institutional Review Board Statement: Ethical review and approval were waived for this study, as no interventions were included in the data collection about humans.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Kwon:, Y. The process of selecting the planned site for an administrative complex city. *J. Geogr. Soc. Korea* **2005**, *39*, 193–203.
2. Hong, S.; Kweon, I. Evaluation on the construction of multifunctional administrative city (Phase 1) through importance–performance analysis (IPA). *J. Korean Reg. Dev. Assoc.* **2017**, *29*, 1–16.
3. Song, Y.-I.; Rhim, J.-H. Economic self-sufficiency criteria for new town planning by network characteristics. *LHI J. Land Hous. Urban Aff.* **2016**, *7*, 251–259. [CrossRef]
4. Son, J. Self-sufficiency and development possibility of new types of cities. In Proceedings of the KGS Conference, Seoul, Korea, 1 December 2007; Volume 2007, pp. 22–53.
5. Lee, S.; Lee, H. Megacity Strategies for Both the Metropolitan Area and Provinces. *Issue Diagn.* **2020**, *433*, 1–26.
6. Jo, M.R. Multifunctional administrative city promotion plan. *Open Chungnam Rev.* **2010**, *52*, 8–26.

7. Elmouelhi, H. New administrative Capital-Cairo: Power, Urban Development and Social Injustice—The Official Egyptian Model of Neoliberalism. In *Neoliberale Urbanisierung*; Al-Hamarneh, A., Margraff, J., Scharfenort, N., Eds.; Transcript Verlag: Bielefeld, Germany, 2019; pp. 215–254.
8. Rachmawati, R.; Haryono, E.; Rohmah, A.A. Developing smart city in the new capital of Indonesia. In Proceedings of the IEEE International Smart Cities Conference (ISC2), Manchester, UK, 7–10 September 2021; Volume 2021, pp. 1–7.
9. Kwon, Y.; Si, S. Sejong Si (City): Are TOD and TND models effective in planning Korea's new capital? *Cities* **2015**, *42*, 242–257. [CrossRef]
10. Abdelrahman Hussein, A.A.A.; Pollock, E. Sustainable Development approaches in Egypt. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *297*, 012027. [CrossRef]
11. Ghalib, H.; El-Khorazaty, M.T.; Serag, Y. New capital cities as tools of development and nation-building: Review of Astana and Egypt's new administrative capital city. *Ain Shams Eng. J.* **2021**, *12*, 3405–3409. [CrossRef]
12. Gomeseria, R.V. Challenge of Environmental Advocacy in Construction Industry. *Durreesamin J.* **2018**, *4*, 1–8.
13. Moser, S. Putrajaya: Malaysia's new federal administrative capital. *Cities* **2010**, *27*, 285–297. [CrossRef]
14. Almdhun, H.M.; Mallak, S.K.; Aburas, M.M.; Md Said, M.A.; Ghadiri, S.M. Measuring and predicting urban growth patterns and trends in Putrajaya, Malaysia. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *169*, 012114. [CrossRef]
15. Shimamura, T.; Mizunoya, T. Sustainability prediction model for capital city relocation in Indonesia based on inclusive wealth and system dynamics. *Sustainability* **2020**, *12*, 4336. [CrossRef]
16. Zou, Y.; Zhao, W. Making a new area in Xiong'an: Incentives and challenges of China's "Millennium Plan". *Geoforum* **2018**, *88*, 45–48. [CrossRef]
17. Zou, Y.; Chen, Z.; Zhong, N.; Zhao, W. Urban planning as a way to pursue quality-oriented urbanization: Anatomy of the urban planning of Xiong'an New Area, China. *J. Urban Aff.* **2021**, 1–16. [CrossRef]
18. Im, C.-W.; Kim, C.-G. A study on the sustainability of the new towns in Tokyo metropolitan area: Focusing on self-sufficiency and social-mix. *LHI J. Land Hous. Urban Aff.* **2011**, *2*, 19–33. [CrossRef]
19. Byun, B.S.; Joo, Y.J. Establishment of indicators and evaluation for sustainable new city land use. *J. Korea Plan. Assoc.* **2000**, *35*, 69–84.
20. Byeon, C.H. Improvement of housing lot development for increasing self-sufficiency of the new town. *Hous. Stud.* **2005**, *13*, 175–208.
21. Seo, J.; Lee, J. Analyzing the Allocation Criterion of Land Use to Enhance Self-sufficiency of New Towns. *Korea Spat. Plan. Rev.* **2011**, *69*, 157–177.
22. Jeong, D.; Kim, H. Analyzing the levels of self-containment and centrality of the five first-period new towns built in the Seoul metropolitan area. *J. Korean Urban Geogr. Soc.* **2010**, *13*, 103–116.
23. Lee, J.H.; Ban, J.H.; Song, T.J.; Hong, S. An analysis of economic self-sufficiency in new towns using commuter traffic-focused on new towns built for balanced development. *J. Archit. Inst. Korea* **2022**, *38*, 117–128.
24. Lee, S.; Joo, M.; Ha, J. An analysis of changes in commuting characteristics and urban spatial structure of the first generation new towns in the Seoul metropolitan area (1996–2010): Focused on self-containment and centrality. *J. Korea Plan. Assoc.* **2015**, *50*, 5–23. [CrossRef]
25. Kim, T.H.; Kim, H.-S. A study on the change of self-sufficiency in Seoul metropolitan area–supply-side analysis of the data between 1995 and 2010. *J. Korea Plan. Assoc.* **2014**, *49*, 23–39. [CrossRef]
26. Cho, M. A Study on Evaluation of the Self-Sufficiency of Cities Using Work Trips in the Seoul Metropolitan Region. Master's thesis, Hanyang University, Seoul, Korea, 2011.
27. Horner, M.W.; Mefford, J.N. Investigating urban spatial mismatch using job–housing indicators to model home–work separation. *Environ. Plan. A* **2007**, *39*, 1420–1440. [CrossRef]
28. Gordon, P.; Kumar, A.; Richardson, H.W. The influence of metropolitan spatial structure on commuting time. *J. Urban Econ.* **1989**, *26*, 138–151. [CrossRef]
29. Ta, N.; Chai, Y.; Zhang, Y.; Sun, D. Understanding job-housing relationship and commuting pattern in Chinese cities: Past, present and future. *Transp. Res. D* **2017**, *52*, 562–573. [CrossRef]
30. Zhang, Z. A study on urban spatial structure in the context of the jobs-housing balance: A case of Suzhou, China. In *Smart Growth Sustainable Development*; Springer: Cham, Switzerland, 2017; pp. 73–85.
31. Zhao, P.; Lü, B.; de Roo, G. Urban expansion and transportation: The impact of urban form on commuting patterns on the city fringe of Beijing. *Environ. Plan. A* **2010**, *42*, 2467–2486. [CrossRef]
32. Yoon, J.J. A Study on the Direction of the Third Phase New Town Development in Seoul Metropolitan Area through expert survey method. *LHI J. Land Hous. Urban Aff.* **2019**, *10*, 43–55.
33. Kim, S.; Lee, D.; Moon, S.; Byeon, S.; Kim, J. Evaluation and directions of New Town development policies in the Seoul metropolitan area. *KRIHS Policy Brief* **2021**, *822*, 1–8.
34. Jo, P.; Jeong, W. Multifunctional administrative city construction process and completion task. *KRIHS* **2021**, *471*, 18–23.
35. Oh, S.; Lee, W. A time series analysis of urban structural change in Pyeongtaek City. *J. Korea Plan. Assoc.* **2012**, *47*, 33–44.
36. An, Y.; Lee, S. A study on the construction of spatial database and the characteristics of relocation for the relocated firms by industrial types. *J. Korea Plan. Assoc.* **2014**, *49*, 17–28. [CrossRef]
37. Ban, J. Income redistribution and poverty alleviation effect by housing type. *Lab. Rev.* **2009**, *54*, 67–83.
38. Kim, H.-K.; Hong, W.-H.; Kim, Y.-C. A new estimation model of MSW generation based on household composition-types and floor area in apartment house. *J. Korean Hous. Assoc.* **2012**, *23*, 41–49.

39. Jang, S.; Yi, C. Imbalance between local commuting accessibility and residential locations of households by income class in the Seoul Metropolitan Area. *Cities* **2021**, *109*, 103011. [CrossRef]
40. Matkan, A.A.; Shahri, M.; Mirzaie, M. Bivariate Moran's I and LISA to explore the crash risky locations in urban areas. In Proceedings of the Conference of Network-Association of European Researchers on Urbanisation in the South, Enschede, The Netherlands, September 2013; pp. 12–14.
41. Anselin, L. *GeoDa 0.9 User's Guide. Spatial Analysis Laboratory*; University of Illinois: Urbana-Champaign, IL, USA, 2003.
42. Yang, D.; Ye, C.; Wang, X.; Lu, D.; Xu, J.; Yang, H. Global distribution and evolvement of urbanization and PM2.5 (1998–2015). *Atmos. Environ.* **2018**, *182*, 171–178. [CrossRef]
43. Lee, C. Analysis of the influence factors of agglomeration by industry in eup, myeon, and dong in the metropolitan area using spatial metrology model. *GRI Res. J.* **2016**, *18*, 1–33.
44. Cho, M.; Song, J. Regional characteristics of migration inflow of one-person households in the Seoul metropolitan area: Focusing on Migration Motivation between Job and Housing. *J. Korea Plan. Assoc.* **2020**, *55*, 70–84. [CrossRef]

sustainability

MDPI

Article

Roadmap for Future Mobility Development Supporting Bangkok Urban Living in 2030

Nathasit Gerdsri [1,*], Komkrit Sivara [2], Chayoot Chatunawarat [2], Somnuek Jaroonjitsathian [2] and Krissada Tundulyasaree [1,3]

1 College of Management, Mahidol University, Bangkok 10400, Thailand; k.tundulyasaree@tue.nl
2 Technology and Engineering Business Unit, PTT (Public) Co., Ltd., Ayutthaya 13170, Thailand; komkrit.s@pttplc.com (K.S.); chayoot.c@pttplc.com (C.C.); somnuek.j@pttplc.com (S.J.)
3 Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands
* Correspondence: nathasit.ger@mahidol.ac.th

Abstract: This study demonstrates how to develop a roadmap guiding a company's technology and innovation for future mobility by integrating the results from foresight and scenario analysis. The key drivers and trends influencing the urban development of Bangkok in the next 10 years were determined through the environmental scanning approach. The participatory approach was then applied to get stakeholders involved to enrich scenario thinking. The participants were invited from many social networks involved in city development in Bangkok. The scenario development was framed according to livability and city infrastructure development. As mobility is always the key issue for big cities, future development for mobility can be addressed in terms of different aspects such as filling stations, clean fuel, public transportation, shared mobility platforms, or last-mile traveling. A roadmap to guide future development of shared mobility was developed based on the analysis of the particular company in the case study.

Keywords: future mobility; shared mobility; future urban living; foresight; scenario development; roadmapping; Bangkok

check for updates

Citation: Gerdsri, N.; Sivara, K.; Chatunawarat, C.; Jaroonjitsathian, S.; Tundulyasaree, K. Roadmap for Future Mobility Development Supporting Bangkok Urban Living in 2030. *Sustainability* **2022**, *14*, 9296. https://doi.org/10.3390/su14159296

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 10 June 2022
Accepted: 25 July 2022
Published: 29 July 2022

1. Introduction

In the age of disruption, changes in society, the economy, law, energy, environment, infrastructure, and technology are swift and unpredictable. In preparing to cope with changes, foresight and scenario development are important tools to understand future trends and build scenarios to illustrate the possibilities. Urbanization has been a global trend in the past decade and has led to new challenges for city development. According to the recent world urbanization report [1], 55% of the world population lives in urban areas, compared to only 30% in 1950. This number is expected to rise to 68% by 2050. This projected number of urban residents is based on the continuity of the urbanization trend and the growth of the world population. Most of the growth is forecasted to be in Asian and African cities. With this growing number of residents, increased demand for various urban resources and the need for proper facility development are inevitable.

To cope with future urban challenges, cities need to create a plan to guide future development. City planners usually face two scenarios: building from scratch or improving on the existing area, with the latter occurring most often. Urban planning also covers many aspects of the city, such as infrastructure, land use, economic development, neighborhoods, and housing. Recently, several researchers have worked on finding new ways to support the future development of cities by addressing issues like mobility, the sharing economy, food sustainability, and active community engagement. To face the future, cities first need to specify the development goal. To set a goal, planners need to study the existing system and obtain the requirements from city stakeholders such as residents and government officials. With agreeable goals, a strategy for development can be defined.

The Bangkok Foresight 2030 project was initiated for two reasons. The first was to exchange ideas and increase social awareness by exploring the possible future of Bangkok in the next 10 years. The second was to provide input to business executives and managers based on the results of foresight and scenario analysis to plan their technological adaptation and future innovation by incorporating such plans into their corporate roadmap.

This paper is organized into six sections. After the introduction in Section 1, Section 2 reviews the latest literature on urban development as well as the applications of foresight, scenario analysis, and roadmapping. Section 3 addresses the methodology and approach used in this study. Section 4 presents the analysis results, ranging from environmental scanning to scenario development; four scenarios are described, along with their illustrations. Section 5 explains how business executives and managers can use the results to develop their corporate roadmap. Finally, Section 6 addresses the conclusion and areas for future development.

2. Literature Review

2.1. Smart Cities: The Forthcoming Trend for Cities

Realizing the coming trend of urbanization, various cities are adopting the smart city concept to tackle challenges and plan for the future. Smart cities can facilitate other urban development goals, such as reducing greenhouse gases and energy consumption or increasing sustainability [2]. The future development of cities is targeted to increase their economic competitiveness and enhance residents' quality of life [3]. For example, Amsterdam and Rotterdam aim to cut down their carbon emissions up to 50% by 2025 [4]. They are planning to increase the share of electric vehicles and bicycles significantly by providing proper mobility infrastructure such as charging stations and cycling lanes. Barcelona has adopted new technologies and fostered collaboration between city stakeholders to transform the city by focusing on the number of jobs, social housing, larger green areas, and better public services [5]. Although these cities are achieving a positive effect using the smart city concept, they seem to define the concept differently.

Academia, governments, and industry have proposed more than a hundred definitions of a "smart city" [6]. Apart from the number of definitions, the concept has been applied to two distinct domains: hard and soft [7]. The hard domain refers to buildings, energy grids, and logistics, whereas the soft domain refers to education, culture, and policy interventions. According to a bibliometric analysis of the smart city literature from 1992 to 2012, Mora et al. highlighted that the overlap between the proposed definitions hinders scholars from achieving a consensus about the concept [8]. From an analysis of more than 100 definitions, Kondepudi defined the smart city concept as follows: "A smart sustainable city (SSC) is an innovative city that uses information and communication technologies (ICTs) and other means to improve quality of life, efficiency of urban operation and services, and competitiveness, while ensuring that it meets the needs of present and future generations with respect to economic, social and environmental aspects" [6]. The emergence of new technologies such as cloud computing, big data, and the Internet of Things has provided a foundation for smart cities, because they provide a connection between elements in the city and create a large data set for the city [9]. Lee et al. have suggested that the concept of smart cities could solve existing urban issues such as the unavailability of public service, traffic problems, and environmental pollution [3]. Moreover, information and communication generated through smart cities could also change residents' behavior [10]. For example, with sensors on buses, arrival times could become more predictable, so passengers may find commuting on public transport more attractive. Apart from providing quality services, smart cities also contribute to urban development. Anthopoulos and Vakali proposed the alignment of the smart city concept and urban development [11], whereas Albino et al. emphasized that the concept should focus on people's and community needs [7]. A smart city is similar, as a concept, to the digital and intelligent city concepts, which involve the adoption of ICT to deal with urban issues [12]. A digital city merges urban data and constructs a virtual space as a platform for communities to exchange ideas [13]. The digital city only focuses on the technological side, whereas smart cities extend the concept to include human and governance perspectives. An intelligent city is also equipped with physical and ICT infrastructure [14]. Yin et al. indicated that the terms "intelligent" and "smart" city have usually been used interchangeably in the literature [9]. However, the latter seems to be a more popular concept because the word "smart" is more user friendly and less political [12].

2.2. Key Elements for City Development

Giffinger et al. identified the commonality among definitions and proposed key elements cities must take into account for the development of a smart city [15]; these include smart mobility, a smart economy, smart living, smart governance, smart people, and a smart environment. Each characteristic is described by a different set of factors, depending on the context and the development stages. For example, the brownfield smart city project is likely to focus on optimizing the current infrastructure rather than building from scratch. The characteristics and their corresponding factors are summarized as follows.

First, a smart economy focuses on enhancing the economic competitiveness by factors such as productivity, entrepreneurship, or innovation. Second, the concept of 'smart people' tackles both the individual and social interaction aspects of the citizens. The former includes the level of education and learning skills, whereas the latter deals with flexibility, social and ethnic plurality, and open-mindedness. Third, citizens' life quality is addressed in smart living from different perspectives such as healthcare and housing conditions. Fourth, smart governance aims to promote citizen participation in politics and transparent administration. Fifth, smart mobility includes both transportation for the area's physical connection and ICT for digital connections. Finally, the smart environment deals with the natural resources and how to handle them sustainably.

Infrastructure is another important element of the smart city. As a foundation of the city, the infrastructure has a direct influence on achieving the smart city goal. The UNCTAD Secretariat categorized smart infrastructure into two groups: physical and digital [16]. Physical infrastructure consists of six sub-elements: buildings, mobility and transport, energy, water management, waste management, and healthcare. All these sub-elements are crucial systems necessary for living. Digital infrastructure refers to the data platform facilitating the accumulation of data. These data may then be analyzed and used by related parties such as city stakeholders to improve citizens' quality of life or to further support the integration of physical infrastructure via the information-sharing platform.

2.3. Challenges and Difficulties in Smart City Development

The development and implementation of a smart city always face various challenges, including technological aspects, infrastructure, legislation, and funding and stakeholders' expectation. Without accurately and comprehensively addressing those challenges, the development would run into problems. For example, an inadequate approach to managing various types of new mobility services would lead to safety issues for users [17,18]. Improper health measures for shared mobility would lead to health concerns during the COVID-19 pandemic [19]. In addition, insecure data-sharing systems would lead to privacy and security issues for shared mobility and Mobility as a Service (MaaS) [20].

Financial and citizen-related issues are common challenges in the implementation of smart city projects, which usually require a significant level of investment to build or improve city infrastructure. For example, China and India have allocated several billion USD to transition to smart cities [21]. Given the huge investment required, a local municipality alone cannot fund a smart city project, so they need to find other sources of funding. For example, a crowdfunding platform was a success in Estonia to build the PARKI science and industrial parks for promoting green technology [16]. This project not only makes money, but the citizens are also more "invested" in the project as they are participating in and contributing to transforming their city.

The stakeholders involved in city development consist of four sectors: governmental and public, university academics and research and innovation, IT and the private sector, and civil society [22]. Each sector has different functions and needs, so collaboration is key to achieving project success. For example, university academics and their research teams can come up with a sophisticated smart city solution, but they may not be able to satisfy the needs of citizens, and their plans may incur a cost too high for the city to afford. Each city also poses different challenges, as they have different starting conditions, available resources, and citizen willingness [23]. The solution in a city thus cannot be immediately adopted in other cities: There are no one-size-fits-all solutions. However, we could still learn from the challenges faced by other smart city projects and how they deal with those challenges. Apart from participation, lack of citizen skills can be a hurdle to smart city projects. Curriculum reform or an online platform such as massive open online courses (MOOCs) may be a potential solution. Should both financial and citizen skill issues be sorted out, the smart city project can be initiated properly.

2.4. Foresight

Technology foresight has been introduced and widely practiced to help organizations address their long-term planning. The term "technology foresight" emerged from technology forecasting, with the main purpose of visualizing the future [24–26]. Technology foresight was first introduced to describe the techniques, mechanisms, and procedures for strategically identifying the potential areas of basic research [27,28]. Technology foresight was applied as a process to understand the forces shaping the long-term future using qualitative and quantitative means for policy formulation, planning, and decision-making [29]. Coates concluded that foresight prepares us to meet the needs and opportunities of the future [30].

Technology foresight involves systematic attempts to look into the longer-term future of science, technology, economics, the environment, and society with a view toward identifying emerging technologies to yield the greatest economic and social benefits [31]. The objective of technology foresight is usually dependent on the individual group of practitioners, such as the government and corporations at both the national and regional levels. However, the most common objectives of technology foresight exercises are to set priorities in science and technology development. The decision results can then be revealed through the allocation of resources, including research funding, scientific instrumentation, and future requirements for trained researchers.

2.5. Recent Foresight Studies on Future Cities

In 2013, the UK Government Office for Science conducted a foresight study to articulate the long-term views on how people will live, work, and interact in more than 20 UK cities in 2065 [32]. The study collected evidence from working papers and essays, as well as running future workshops and visiting, supporting, and working with cities of various types and sizes across the country. The results offer insights for policymakers at the national level, as well as to local government officials and partners, for understanding the future of cities and in what direction research could most usefully be focused in the future.

In 2014, Shell Corporation released the scenario study "New Lens on Future Cities", which aimed to clarify the differences in how cities are built and operated to make the right choices for building sustainable cities [4]. Shell's study projected that urban populations are growing and around three out of every four of people on earth will live in cities by 2050. As cities expand, pressure on vital resources such as energy, water, and food increases. According to Shell's scenarios, cities can be grouped into six categories or archetypes. Each category was analyzed to create scenarios for how individual cities could evolve and become more efficient.

Tewdwr-Jones and Goddard applied foresight to long-term urban planning to address the future opportunities and threats for land use [33]. They also showed that foresight activities can encourage partnership and potentially lead to collaboration among government, universities, the private sector, and citizen stakeholders. Fabbri applied foresight to strategic planning processes to accelerate the development of regional innovation in Tuscany [34]. Similarly, foresight was also applied to understand potential sustainable urban futures for China in 2050 [35]. Recently, there was also a foresight study on the urban mobility of Singapore in 2040, conducted by Zahraei et al. [36]. Two scenarios were presented, called the Shared World and the Virtual World. Transport modes for passengers and freight were highlighted in each scenario, along with possible implications to the individual, society, industry, and government levels. Although there have been several foresight studies related to the future city, Güell and López still raised concerns about the appropriateness of the foresight tools used in cities and the competency of foresight practitioners in understanding the complex and dynamic nature of contemporary cities [37].

2.6. Roadmapping and Extended Applications

Roadmapping is the process for linking a business strategy with a technology strategy. A set of alternatives for meeting certain objectives can be identified along with critical requirements for a set of needs [38]. Roadmaps can address different aspects of a planning problem and subsequently provide a consensus view or vision of the future landscape available to decision-makers [39,40]. The graphical nature of a roadmap visually portrays links and relationships between capabilities and requirements [41,42]. The roadmap document presents a potential future by focusing on specific knowledge from a chosen field [43], and the ability to display and connect time-based knowledge is what makes technology roadmapping (TRM) a powerful method to support strategic planning [44].

Several scholars are continuing to work on developing additional tools to customize and operationalize the roadmapping process. For example, Kostoff et al. proposed a systematic approach to identify disruptive technologies by taking advantage of text-mining [39,45]. Passey et al. studied integration product concept visioning and scenario building with roadmaps [46], and Lizaro and Reger proposed connecting roadmapping and scenarios to plan the coordinated development and deployment of new and existing technologies and applications [47]. Satangput et al. integrated scenario development into roadmapping to help the APEC community identify development areas for future technologies to combat emerging infectious diseases [48]. Lee et al. utilized patent information for new product development by developing a keyword-based TRM approach [49]. Yasunaga et al. studied the utilization of TRM with the structuring of knowledge in R&D management [50]. Gerdsri and Kocaoglu proposed the technique called technology development envelope (TDE), which uses a Delphi survey to identify trends in emerging technologies, and the analytical hierarchy process (AHP) to sequentially evaluate the impact of emerging technologies on a company's objectives over time [51]. Yoon et al. applied morphology analysis to derive promising opportunities for the development of new products or technology by matching their morphology [52]. Chutiwongse and Gerdsri applied roadmapping to help managers guide their organizational development activities and organizational innovation [53]. Gerdsri and Manotungvorapun emphasized the need for assessment and roadmapping to support the development of effective university–industry collaboration (UIC) [54].

3. Methodology, Approach, Engagement, and Workshop Activities

To develop a roadmap guiding technology and innovation development for future urban living in Bangkok, the approach used in this study was designed by integrating three methodologies: foresight, scenario analysis, and roadmapping, as shown in Figure 1 below.

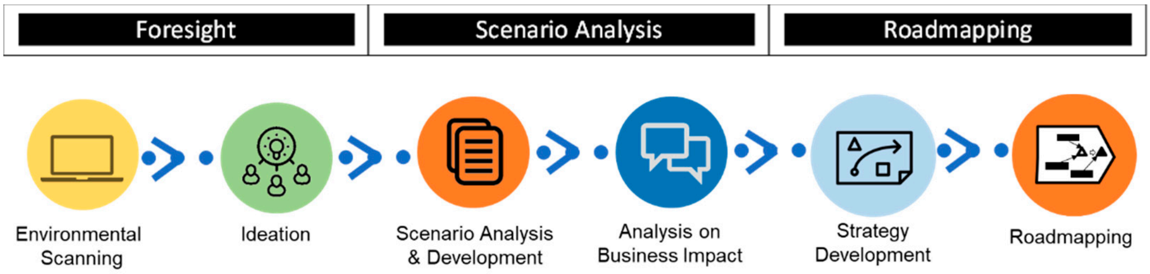


Figure 1. Integrating foresight and scenario planning with the roadmapping approach.

3.1. Methodology

3.1.1. Foresight

Foresight was applied to determine the future trend of key drivers following the environmental scanning approach. This was carried out based on the assumption that an uncertain future is shaped by today's actions. Key experts from different areas influencing the future development of Bangkok were invited to share data and their views.

3.1.2. Scenario Analysis

The scenario analysis was implemented to create a portrait of the future possibilities. Livability and city infrastructure development were the two key dimensions to frame four different scenarios, as shown in Figure 2. Livability represents a community's quality of life, including economic prosperity; social stability and equity; educational opportunities; public health, safety, and food/water access; and culture, entertainment, and recreation possibilities. City infrastructure development means buildings, developments, or improvements to the transportation or telecommunication systems, as well as land and resource use. The description of each scenario is provided below.

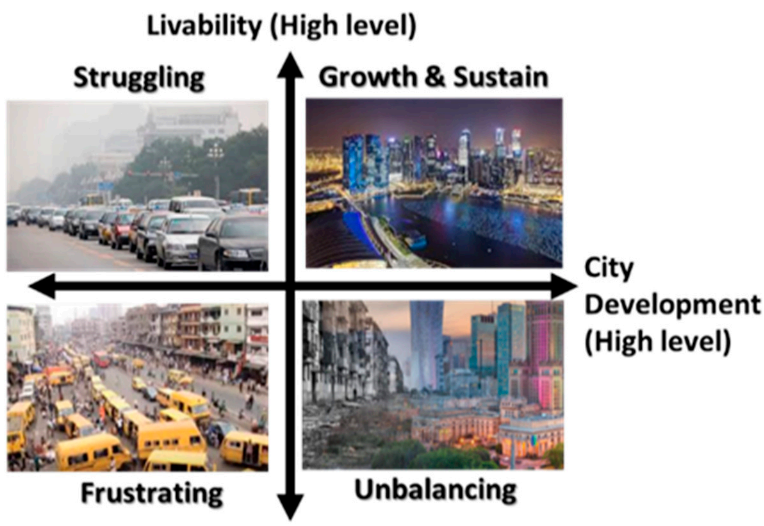


Figure 2. Scenario analysis based on "livability and city infrastructure development."

Scenario 1: Growth and Sustainability (i.e., a high level of infrastructure development and a high level of livability). This scenario represents the balanced development of infrastructure and livability, which leads to growth and sustainable development of the city.

Scenario 2: Struggling (i.e., low level of infrastructure development and an increasing or high level of livability). This scenario focuses on the rise of economic development in the city while the development of infrastructure such as utilities and transportation systems still lags behind or is improperly planned.

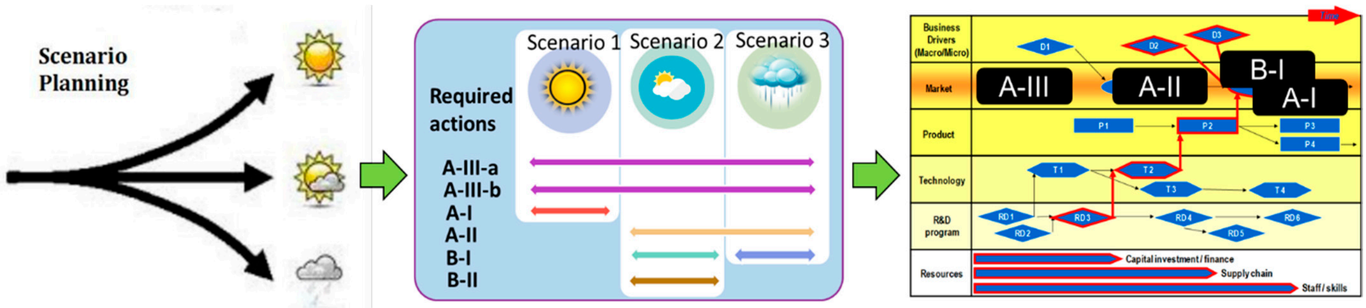
Scenario 3: Frustrating (i.e., low level of infrastructure development and low level of livability). This scenario represents the situation where the economy is under-developed, causing many problems in basic healthcare services and food access, among others. Infrastructure development is also insufficient.

Scenario 4: Unbalancing (i.e., high level of infrastructure development and decreasing or low level of livability). This scenario represents unbalanced development, which can be explained by two phenomena: One is the shrinking of cities or urban depopulation resulting from economic turmoil or political crisis, and the other is the condition in which advanced development happens only in certain areas, so only a limited number of people can benefit from it. This creates a social divide.

In this study, the scenarios were developed by combining both normative and exploratory approaches with a mix of subject experts and wider participants, similar to the mix-methods approach introduced by Star et al. [55].

3.1.3. Roadmapping

Based on the developed scenarios, business executives and managers can analyze the impacts on their business, as well as identify emerging opportunities before developing a strategic plan and roadmap to guide their technological adaptation and future innovation. This strategic plan is often presented as a company roadmap. In constantly changing situations, it is always difficult to specify which scenarios are most likely to occur. Often, executives decide to wait and see before taking action, but this approach is passive and risky. In this study, we tackle these challenges by taking an adaptive approach. The key consideration is to identify what should be done no matter which scenario takes place (i.e., identifying common issues required regardless of which scenario appears), and then setting these actions as part of a short-term plan in the roadmap. Figure 3 shows these actions; preparations should be made for A-IIIa and A-IIIb regardless of whether it is sunny, cloudy, or rainy, and A-III should be set as the short-term target for development on the roadmap.



(**a**) Scenario development (**b**) Scenario/impact analysis (**c**) Roadmapping

Figure 3. Adaptive approach for connecting scenario/impact analysis into roadmapping.

3.2. Step-by-Step Approach

To achieve the objective of this study, the analysis approach was divided into eight steps, as shown in Figure 4 below. Steps 1 and 2 are the project preparation, in which the core project team worked together to review the key literature using foresight/scenario analysis/roadmapping as a part of the methodology, as well as the studies related to future urban living. The core team also set up a framework to identify key factors according to their impacts on the development of cities over the past 10 years as well as their potential impacts on shaping the future development of cities [4]. The core team also specified the required data type and sources, and identified potential experts to be invited to share their insights during the environmental scanning session. Step 3 was set for environmental scanning to comprehend the key trends influencing the future development of urban living in Bangkok. At this stage, the process was referred to as “Foresight”. Step 4 was the scenario analysis and development. Steps 5–8 were the plans to work with business executives and managers to analyze the impact on a business, explore new opportunities, consider alternatives for technological adaptation and innovation, and develop a company roadmap. The processes for these steps are business impact analysis, strategy development, and roadmapping.

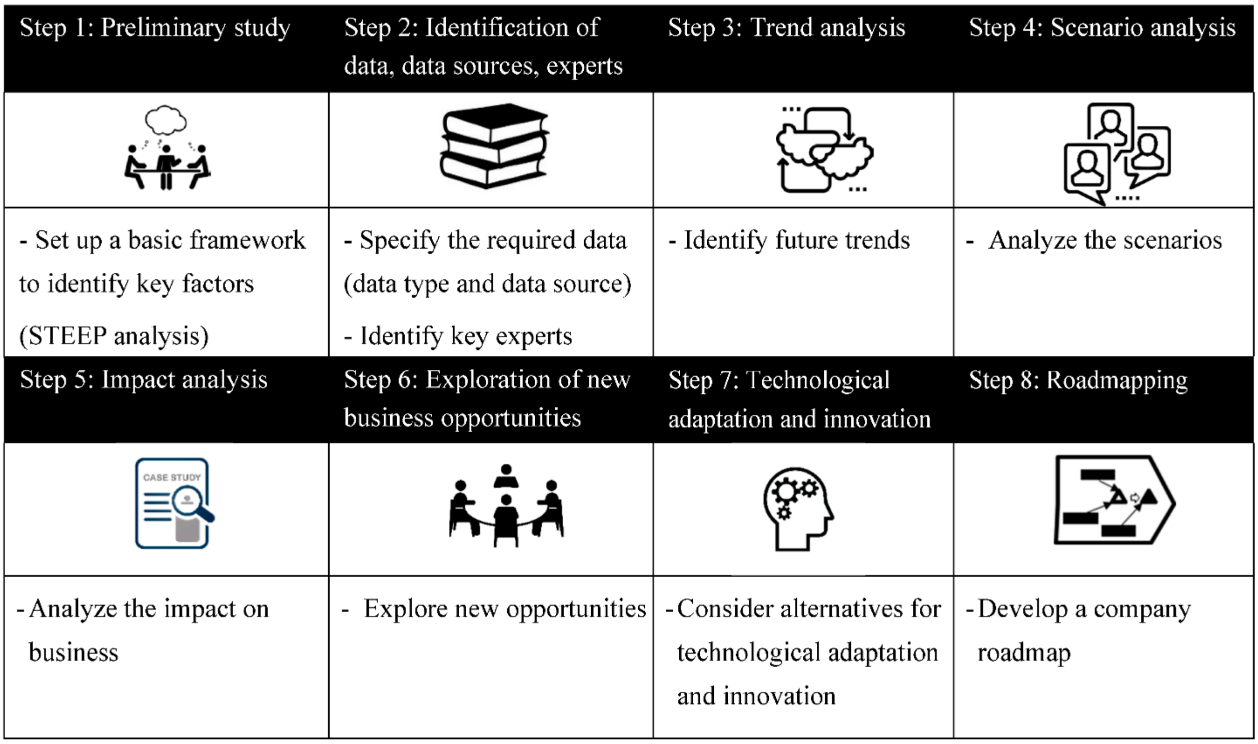


Figure 4. Eight-step approach.

3.3. Engagement

The participatory approach was applied to get stakeholders involved in scenario thinking. The group of participants were invited from many social networks where they were considered opinion leaders. Altogether, there were around 60 people in the age range between 25 and 40 from diverse professional and education backgrounds. Those participants also represented both the public and private sectors. We considered this group to be made up of active citizens. Participants were then arranged in four groups to develop scenarios articulating their future thinking. This process was referred to as scenario analysis and development.

3.4. Workshop Activities

The foresight workshop was organized over the course of two full days. The first day was designed for environmental scanning. The comprehensive summary was provided to the participants by experts in different areas, such as:

- The Minister of Transportation shared the development plan for the mass transit systems in Bangkok;
- The top management of Toyota Corp. (Toyota, Japan) and LINE Corp. (Tokyo, Japan) shared the outlook for communication and mobility technologies (note: LINE Corp. is a Tokyo-based company focusing on the development of mobile applications and Internet services, particularly short message communication and entertainment);
- A university professor expert in urban planning shared their research on demographic data, living and lifestyle, land use and zoning, impacts of rising sea level and environmental issues;
- The lead economist from the Ministry of Finance shared the economic outlook of Bangkok and major cities in the region; and
- A lawyer from a leading international law firm in Bangkok shared their thoughts about the future direction of legal development in Thailand, as well as a comparative study with other countries.

On the morning of the second day, the participants were organized into groups to articulate scenario thinking using the information from the environmental scanning provided the previous day. Key issues for each scenario were identified. In the afternoon, each group worked on scenario development, including storylines and graphic illustrations. At the end of the day, each group shared their scenario and then the group opened up for wide-ranging discussion. After the workshops, the results were also shared with the public. Business executives and managers can use the results from the environmental scanning to analyze the impacts on their business as well as to identify emerging opportunities before developing a strategic plan by following Steps 5–8, as described in Figure 5. For the particular company referred to in this study, a working group of business executives and managers was formed to analyze the impacts of each scenario on the company's business and identify the potential for new business opportunities. The internal process of going through Steps 5–8 took around three months.

4. Results Analysis

In this section, the analysis results are presented according to the proposed approach, linking foresight to scenario analysis and roadmap development. The key findings are also highlighted in each approach.

4.1. Foresight: Environmental Scanning

The details about the key drivers and trends influencing the future development of Bangkok, which were extensively discussed during the workshop, are summarized below.

4.1.1. Infrastructure Development: Bangkok Mass Transit Systems

Thailand aims to employ a railway system including high-speed trains as the main mode of transportation and connecting to all neighboring countries. The urban train lines will be extended to increase coverage to more than 400 km. High-speed rail systems are also under construction to connect Bangkok to different regions and neighboring countries [56].

4.1.2. Technology and Online Society Development

At present, internet usage among people in Thailand is in the world's top rank, where people spend on average of 216 min per day on smart devices, with about one-third using LINE applications. A LINE executive mentioned that the company aims to provide well-connected, fast transactions and tailored services for convenience. Moreover, he also sees the need to transform content into having "everything videologized", because he found that people in Thailand now spend, on average, 40 min on YouTube each day.

4.1.3. Economic Development

Although the current economic growth and unemployment rate are healthy, there is still room for improvement, as there is imbalance in the four important economic drivers: private consumption, public consumption, investment, and export. The concentration of wealth is also increasing. To tackle these issues, the government should prioritize economic development of rural areas to decrease the migration of workers toward big cities. Concurrently, they should also monitor global mega-trends such as cryptocurrency, blockchain, the sharing economy, and green economy to explore potential opportunities. Finally, it is necessary for both the public and private sectors to cooperate to drive the country forward.

4.1.4. Urban Development

Another important factor is quality of life. Due to urbanization, the sense of ownership among urban citizens starts to decline. The urban planning expert stressed that the government should pay more attention to the design of public spaces to serve citizens. He suggested that, in the future, citizens will spend more time in public spaces than in their house (or "small room").

4.1.5. Environment and Energy Planning

Bangkok is facing similar challenges to those of other metropolises, such as the rising number of residents, increased traffic jams, uncertain public safety, and information privacy. These alter the city's energy consumption pattern. Decarbonization is a new trend to tackle urban pollution, and decentralized energy generation is also becoming more popular.

4.1.6. Food

Due to the developed logistics system, people can easily access food delivery services, but the increasing cost of transportation affects the cost of food services. To deal with this issue, new technologies should be considered to optimize food distribution. The application of smart farming can also decrease the cost of production, as it needs fewer resources to produce. Another challenge is to produce enough food for the world population, because the population continues to grow. Protein from insects is a promising new alternative, as it requires fewer production resources compared to meat. As Thailand is well known as a food producer, there are vast opportunities to position Bangkok as a food hub, especially for street food. If the local community can find a way to enhance standards and quality, it would be a unique factor to attract more tourists.

4.1.7. Law

Due to globalization and future challenges, the current law must be amended. The amendment should not only incorporate infrastructure and economic development but also citizens' quality of life. Recently in Thailand, there have been various legal amendment attempts, such as bills related to civil partnership, the elderly fund, and sale with the right to redemption. Another possibility is to amend the existing law or create new measures to suit future innovation. Those measures or laws include the electronic transaction rule, the information privacy bill, the start-up policy, and the digital industry. However, there are gaps for legal improvement to deal with emerging issues such as who is responsible for the issues related to artificial intelligence, and new laws should also facilitate new types of business such as Airbnb or Uber.

4.2. Scenario Analysis

Four different scenarios were developed according to the two key dimensions—livability and city infrastructure development—as shown in Figure 2. The workshop participants were arranged into four groups. Each group was responsible for the development of one scenario using the key information and trends provided by the experts during the environmental scanning session, including information about the future infrastructure development plan, projected population growth, future technologies, economic development, and food and healthcare access. The description and illustration of each scenario are presented below.

4.2.1. Scenario 1: Growth and Sustainability

Public transportation infrastructure was sufficiently developed to be smart and reliable. It can provide optimized routes and seamlessly shared mobility for people with different lifeways. There are limited parking spaces with a high parking fee for personal cars. Old cars are restricted from the road, and the sense of car ownership has declined; eco-friendly cars are more popular. Due to the smaller number of cars, the average travel time has decreased and the road safety level is higher.

Along with the modern power system, a smart grid is used to manage the whole system to optimize loads. The government supports renewable energy and a free-trade market scheme. Time of use for electricity price rates is employed to realistically reflect the true cost of generation. System stability is enhanced by the energy storage system. In addition, organizations focus on energy management to consume energy effectively.

Not every area in Bangkok is urbanized. Some areas still preserve their identity to attract tourists who prefer the traditional and authentic atmosphere. Pollution measures were implemented, and citizens' environmental awareness was enhanced. Moreover, the PM 2.5 issues were tackled sustainably. The government successfully implemented various policies such as the car restriction policy (increasing annual tax, old cars banned from the inner city center), supporting eco-friendly cars, waste sorting, limiting the amount of waste, recycling water in the canal, and reforming the promenade as a public space.

All citizens can afford the smart healthcare system, which is well-adapted to the aging society. Advanced food-processing technology is used to create nutrient-rich sustenance. Street-food services still remain in Bangkok. Various types of fusion food are introduced, along with new technologies to manage food supply chain and logistics. The description of Scenario 1 can be visualized in Figure 5.

Figure 5. Illustration of Scenario 1: Growth and Sustainability.

4.2.2. Scenario 2: Struggling

Although the public transport infrastructure lacks complete connectivity and stability, the system is still affordable to the majority of citizens. This explains why some citizens choose to travel by public transportation and shared mobility, while others keep using their personal cars for convenience. Traffic and road safety remain issues for the city. Renewable energy is employed to a certain extent, but the power system is still not stable enough. Extra electrical generation is needed during certain periods. Because the energy policy is not clear, each electric producer individually generates electricity without cooperation, so the price does not truly reflect the optimal cost of generation, although it is still affordable for the citizens. System stability would be enhanced by the energy storage system. Many organizations focus on energy management to effectively consume energy.

The tourism sector attracts many foreign tourists to Thailand and especially to Bangkok as the transit gateway. Because the transport infrastructure is ill-prepared, tourists cannot rely on public transportation, so they are arranged in small groups or free individual travelers (FIT) to visit the city. Bangkok residents tend to lend their place via Airbnb. Places close to the metro line are more attractive. Due to the need to communicate with foreign people, the people of Thailand need to improve their English skills.

The pollution control measures are not good enough, so the city struggles with PM 2.5 issues. Without proper car regulation, the number of cars is growing. Moreover, the amount of waste is growing. Even with these issues, Bangkok is still attractive enough to foreign tourists, because they can still enjoy the lack of strict rules and street food. There are no environmental measures to handle the current pollution issue. Due to the rising number of tourists, a variety of food is served on the streets. This causes a shift in consumption behavior related to street food. People cannot control their mealtimes because they cannot be sure of their travel time. In 2030, there are limited healthcare choices for citizens, as the system cannot cover the need for all people. Although the price is still affordable, the number of hospitals and staff is not sufficient to serve the growing population. The description of Scenario 1 can be visualized in Figure 6.

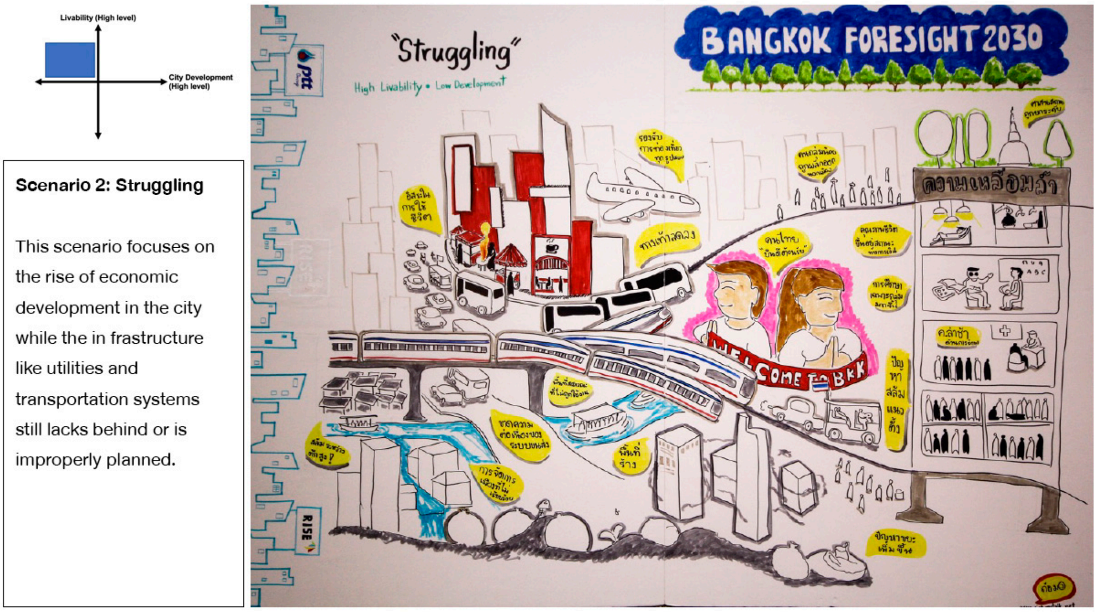


Figure 6. Illustration of Scenario 2: Struggling.

4.2.3. Scenario 3: Frustrating

As the public transport infrastructure lacks complete connectivity and stability, commuters have to pay fares separately for using each system. Total fares are thus too high for the average income. Low-income citizens still need to use public transport because there are no choices. Shared mobility is only affordable to the middle-income range. Although traffic conditions are terrible, personal cars are still an attractive alternative. As the air quality is getting worse, catastrophic warnings to shut schools and work are often used, which greatly affects the economic situation. In addition, due to the congested road conditions and bad-quality cars, road safety has declined, which decreases still further the modal share of active transport modes such as bicycles and scooters. The power system is not stable, so extra electricity generated by private businesses and individuals is needed. However, the demand still cannot be met, as it exceeds generation capacity. Only a few people have started to adopt renewable energy, as there was a lack of government support. Electricity bills are too high, so citizens need to limit their energy consumption. Energy storage and energy management are very rare and only adopted by high-income citizens.

The city center has become a financial district in the daytime and a slum at night. There are many poor and homeless people living on the street. Public transport only operates during the daytime. Some public space has become an exclusive area in which a small group of people share space.

Although the PM 2.5 issue has deteriorated, the government still cannot deal with it. The number of cars is skyrocketing, as there is no ban on old cars. Few people use eco-friendly cars, as the government does not offer any incentive or support. There is no designated area used for waste sorting and dump sites. Waste burning also contributes to the PM 2.5 issue. The community living near rivers continues to dump their waste into the water.

Good-quality food is not affordable for the majority. Some citizens suffer from NCDs (noncommunicable diseases) such as diabetes and hypertension, as the population tends to consume high-sugar foods and drinks. This is in line with the world trend that foods are becoming richer in fat, sugar, salt, and seasoning. Healthcare systems are underdeveloped. There are limited healthcare choices available, and the number of hospitals and available staff is not enough to serve the growing population; the price of medical care is also too high. The description of Scenario 1 can be visualized in Figure 7.

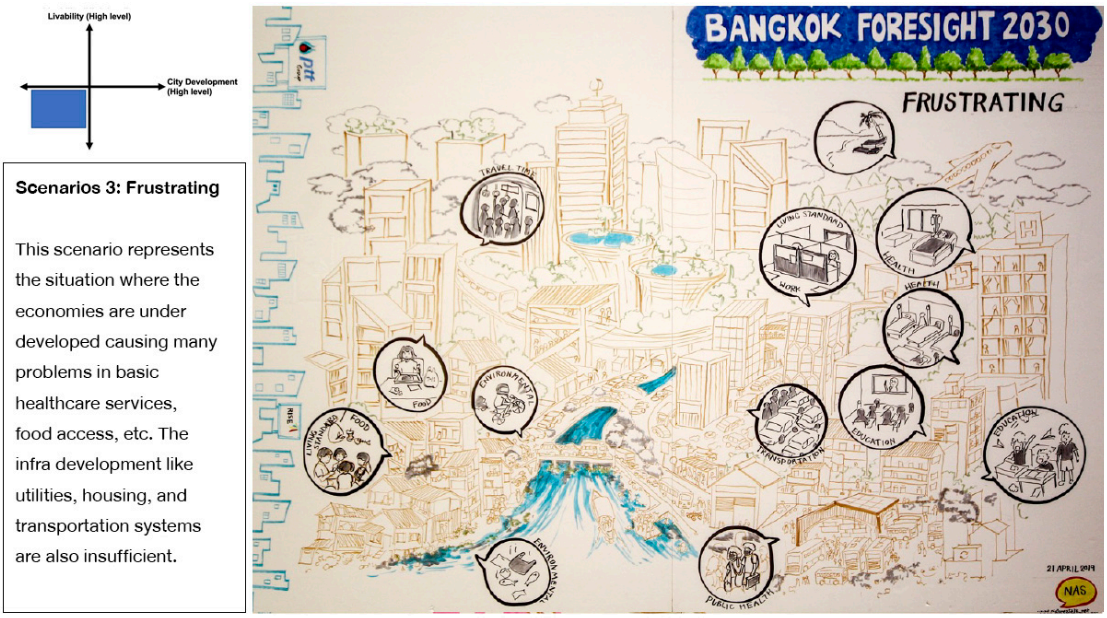


Figure 7. Illustration of Scenario 3: Frustrating.

4.2.4. Scenario 4: Unbalancing

Bangkok lacks well-considered urban planning. The public transportation infrastructure is smart and reliable, but is only available in business-center areas without considering the metro system for the overall city, so only a limited number of people can benefit. This generates denser construction around the metro line, causing property values to increase. Middle-income citizens have to live in the suburban areas, while only the rich, foreigners, and tourists live in the city center.

Citizens have to rely on traditional modes of transportation such as buses and minibuses. Personal cars are still popular among high-income people, and eco-friendly automobiles are widely considered among this group of people to be a luxury. Smart grids are used to manage the whole power system to optimize loads. The government supports renewable energy; however, only high-income people benefit. The free-trade market for energy is employed, and there is no more energy-cost subsidy. The electricity price is charged to reflect the true cost, and the rates are also varied at different times of day. The energy storage system is mainly used for price arbitrage to make a profit.

Middle-income people do not have enough money to afford their own housing, so they become renters. People are aware of environmental problems. The government has implemented a car restriction policy (i.e., rising annual tax, banning old cars from the inner city center) and a taxation policy on recycling water. The cost of living in the city is increasing to the point that middle-income citizens move out to the suburbs and low-income people face difficulty in accessing nutritious foods. Organic products are only affordable by the rich. In 2030, the public healthcare system continues to develop, and the quality is widely acceptable at international standards. However, it cannot serve the whole Bangkok area, as hospitals are concentrated in a certain area and medical costs are quite high for most citizens. The description of Scenario 1 can be visualized in Figure 8.

Figure 8. Illustration of Scenario 4: Unbalancing.

4.3. Roadmap Development

For the particular company in this study, a group of business executives and managers applied the results from the foresight and scenario analysis to plan their technological adaptation and future innovation. The group of executives had interests in the future development of mobility, energy use, and the environment. The opportunities and constraints for future development in those three areas were therefore explored and compared across the four scenarios. Table 1 presents an example of the opportunity analysis on future mobility. The group then began to consider alternatives for technological adaptation and innovation before planning a company roadmap. Due to confidentiality reasons, the company name cannot be revealed.

For future mobility, the group discussed alternatives related to different types of clean fuel, forms of filling stations, modes of public transportation, and platforms for shared mobility, which may be required to serve the development of future mobility in each scenario. Figure 9 shows that the group agreed that the technological adaptation and innovation for a shared mobility platform would be necessary across all four scenarios (represented by a green or solid color-filled box). While comparing the requirements for types of clean fuel, forms of filling stations, and modes of public transportation, there are some variations across the four scenarios (represented by boxes with red diagonal stripes or yellow reverse diagonal stripes). Without wasting time waiting to see which scenario would be the most probable or investing in something that is still not sure, the company can consider taking a proactive approach by setting up the target to drive innovation and R&D activities for the development of a shared mobility platform, because they share something in common across all four scenarios.

The group then continued to identify the key components for the development of a shared mobility platform, including mode (car, bike, scooter), reservation and scheduling management system (year, month, day, time, duration, trip), route (roundtrip, one-way, multiple stops), ownership management system (rental store, P2P), ride control (self-drive, passenger(s), autonomous), and operating system, data management, and user interface. These components are represented in a roadmap as the areas for development. The technological requirements for the future development of each key component were defined, along with the assessment of the company's current capabilities. As a result, the gaps for technological development were determined so R&D projects could be set along with resource planning. If those projects fall within the company's capabilities, then they were planned as an internal project; if not, then the group aimed to work with outside partners. Besides technological development, the company also needs to prepare to educate the market as well as to engage in the policy development [57,58]. Figure 10 shows the roadmap structure designed to guide the development of a shared mobility platform for this company.

Table 1. Opportunity analysis on future mobility across the four scenarios.

	Future Mobility
Scenario 1: Growth and Sustainability	➢ The public transport infrastructure is smart and reliable. It provides a variety of optimized routes and seamlessly shared mobility suitable for every group of citizens. ➢ There are limited parking spaces and with high parking fees. Old cars are banned from the road. The number of drivers and sense of ownership are thus declining. This also decreases air pollution and the number of road accidents. ➢ More and more people are starting to use eco-friendly cars due to their environmental concerns.
Scenario 2: Struggling	➢ The infrastructure for public transportation lacks good connectivity and stability, but the system is still affordable for most of citizens, so some citizens choose to travel by public transportation and shared mobility. ➢ Many keep using their personal cars for convenience. ➢ The government has failed to capitalize on the popularity of the eco-friendly car trend, so air pollution is getting worse. Traffic conditions and road traffic fatality rates are roughly the same as in the present situation.
Scenario 3: Frustrating	➢ The infrastructure for public transportation lacks good connectivity and stability. The ticket prices are too high compared to the average income, but low-income citizens still rely on it because shared mobility is only affordable to the middle class. ➢ Many keep using their personal cars, but they face terrible traffic and spend long hours on the road. This worsens the air quality and has an impact on health. ➢ Due to the poor air quality, there are frequent orders to shut schools and work, which can worsen the economic condition. ➢ Due to the congested roads and poor-quality vehicles, road safety is worse. This also decreases the modal share for active transport modes such as bicycles and scooters.
Scenario 4: Unbalancing	➢ The infrastructure for public transportation is smart and reliable. It provides a variety of routes, but they are not optimally suitable for all citizens. ➢ Most people still rely on traditional modes of public transportation, such as buses and minibuses, because the new system is unaffordable. ➢ Due to the failed car restriction policy, many people keep using their personal cars. ➢ Eco-friendly vehicles are considered luxury items among high-income earners. ➢] From the emissions of all transport modes, air quality is worse, which can lead to serious health problems. ➢ The number of road accidents has decreased, however, because there are fewer cars on the road.

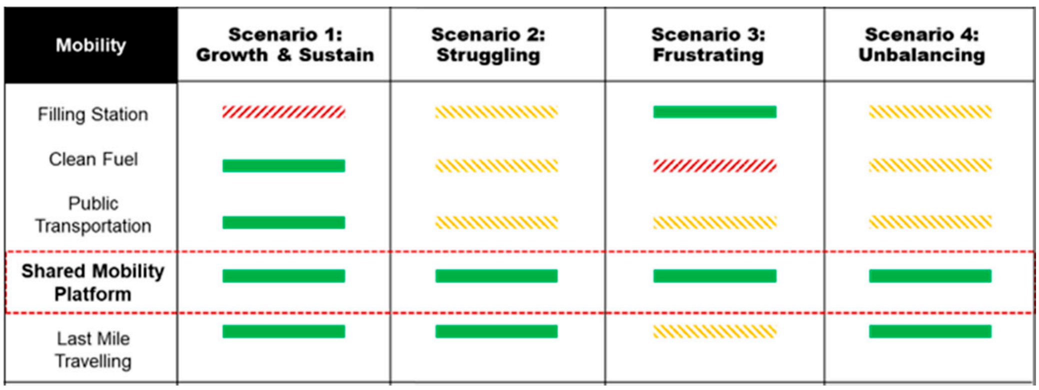


Figure 9. Common requirements across the four scenarios for shared mobility platform (note: if an alternative is strongly required, it represents as a green or solid color-filled box. If "somewhat required", it is represented as a yellow reverse diagonal strip. If "limited required", it is represented as a red diagonal strip).

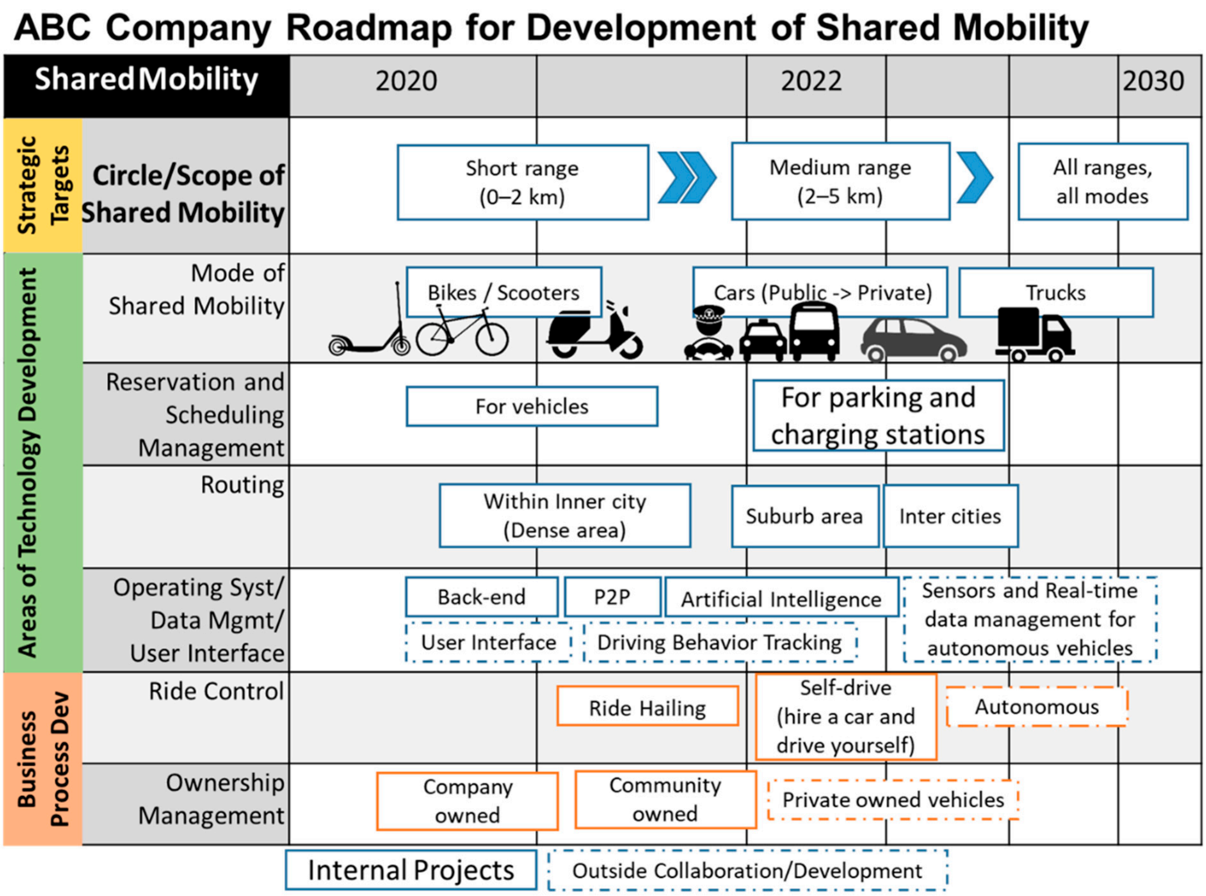


Figure 10. Structure of the company roadmap guiding the development of a shared mobility platform (note: the items with a solid-line box represent the development that will be done by internal teams, whereas the items with a dashed-line box represent the development required by collaborations with outside partners.

5. Discussion

This study reflected that the future development of Bangkok continues to face various challenges. As the results of foresight and scenario planning, the possibilities for how Bangkok could be developed in the future were shown, ranging from the optimistic viewpoint—as presented in Scenario 1: Growth and Sustainability—to the pessimistic one as presented in Scenario 3: Frustrating. The development of four possible scenarios—Growth and Sustainability, Struggling, Frustrating, and Unbalancing, along with imaginative explanation, provided insights and helped the key stakeholders to envision what is likely to come.

The study also showed that mobility is always a key issue for big cities. The future development of mobility can be addressed in different ways, such as through filling stations, clean fuel, public transportation, shared mobility platforms, or last-mile traveling.

The contributions of this study are directed to the community of urban planners and policymakers involved in urban planning, as they can look into the details of each scenario and try to come up with policies to promote preferable conditions, as well as policies to prohibit undesirable conditions such as pollution, social divide, weak economy, or poor infrastructure development.

In addition, the contributions of this study to the case company are directed to the future development of its shared mobility products and services. As follow-up steps, the case company continued to finalize the future characteristics of shared mobility according to customer needs and expectations, and then develop a product-technology roadmap to guide future development of shared mobility solutions in detail. The development teams from both inside and outside of the company were gathered to handle different parts of the project. The core team in charge of the overall development of future mobility solutions worked closely with the development teams to make sure that all activities would be completed on time as planned.

6. Conclusions

This foresight study began with environmental scanning of various drivers of change that could influence long-term urban living and the production of visions of future possibilities. The scenarios were created to shed light on how alternative futures for urban living could unfold as the consequence of different policy directions and the development of particular technologies. The scenario planning was conducted through a series of workshops with the engagement of active citizens from various backgrounds. Storylines and illustrations were also provided to present each scenario. The study also demonstrated the usability of foresight and scenario analysis to develop a roadmap guiding a company's technology and innovation development for future mobility.

Author Contributions: Conceptualization, N.G.; methodology, N.G.; formal analysis, N.G., K.S., C.C., S.J. and K.T.; writing—original draft preparation, N.G. and K.T.; writing—review and editing, N.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived in view that subjects cannot be directly or indirectly identified. Not applicable.

Informed Consent Statement: The study did not affect the rights and welfare of the subjects. Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United Nations, Department of Economic and Social Affairs & Population Division. *World Urbanization Prospects: Highlights (St/Esa/Ser.A/421)*; United Nations Publisher: New York, NY, USA, 2019.
2. Sikora-Fernandez, D.; Stawasz, D. The concept of smart city in the theory and practice of urban development management. *Rom. J. Reg. Sci.* **2016**, *10*, 86–99.
3. Lee, J.H.; Hancock, M.G.; Hu, M.-C. Towards an effective framework for building smart cities: Lessons from Seoul and San Francisco. *Technol. Forecast. Soc. Chang.* **2014**, *89*, 80–99. [CrossRef]
4. Shell Center for Livable Cities. *New Lenses on Future Cities. New Lens Scenarios*; Centre for Livable Cities: Singapore, 2014.
5. Bakıcı, T.; Almirall, E.; Wareham, J. A smart city initiative: The case of Barcelona. *J. Knowl. Econ.* **2012**, *4*, 135–148. [CrossRef]
6. Kondepudi, S.N. *Smart Sustainable Cities: An Analysis of Definitions Series of FG-SSC Technical Reports/Specifications*; International Telecommunication Union: Geneva, Switzerland, 2014.
7. Albino, V.; Berardi, U.; Dangelico, R.M. Smart cities: Definitions, dimensions, performance and initiatives. *J. Urban Technol.* **2015**, *22*, 3–21. [CrossRef]
8. Mora, L.; Bolici, R.; Deakin, M. The first two decades of smart-city research: A bibliometric analysis. *J. Urban Technol.* **2017**, *24*, 3–27. [CrossRef]
9. Yin, C.; Xiong, Z.; Chen, H.; Wang, J.; Cooper, D.; David, B. A literature survey on smart cities. *Sci. China Inf. Sci.* **2015**, *58*, 1–18. [CrossRef]
10. Harrison, C.; Eckman, B.; Hamilton, R.; Hartswick, P.; Kalagnanam, J.; Paraszczak, J.; Williams, P. Foundations for smarter cities. *IBM J. Res. Dev.* **2010**, *54*, 1–16. [CrossRef]
11. Anthopoulos, L.G.; Vakali, A. Urban planning and smart cities: Interrelations and reciprocities. In *The Future Internet*; Álvarez, F., Cleary, F., Daras, P., Domingue, J., Galis, A., Garcia, A., Gavras, A., Karnourskos, S., Krco, S., Li, M.-S., et al., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 178–189.
12. Eremia, M.; Toma, L.; Sanduleac, M. The smart city concept in the 21st century. *Procedia Eng.* **2017**, *181*, 12–19. [CrossRef]
13. D'auria, A.; Tregua, M.; Bifulco, F. Digital city towards smart city: A theoretical overview. In Proceedings of the 2nd International Virtual Scientific Conference, Virtual. 9–13 June 2014.
14. Malek, J.A. Informative global community development index of informative smart city. In Proceedings of the 8th WSEAS International Conference on Education and Educational Technology, Genova, Italy, 17–19 October 2009; WSEAS Press: Athens, Greece, 2009; pp. 17–19.
15. Giffinger, R.; Fertner, C.; Kramar, H.; Kalasek, P.; Pichler-Milanović, N.; Meijers, E. *Smart Cities: Ranking of European Medium-Sized Cities*; Vienna University of Technology, Centre of Regional Science: Vienna, Austria, 2007.
16. UNCTAD Secretariat. *Issue Papers on Smart Cities and Infrastructure*; United Nations Economic and Social Council: New York, NY, USA, 2016.

17. Bozzi, A.D.; Aguilera, A. Shared E-Scooters: A Review of Uses, Health and Environmental Impacts, and Policy Implications of a New Micro-Mobility Service. *Sustainability* **2021**, *13*, 8676. [CrossRef]
18. Turoń, K.; Czech, P. The Concept of Rules and Recommendations for Riding Shared and Private E-Scooters in the Road Network in the Light of Global Problems. In *Modern Traffic Engineering in the System Approach to the Development of Traffic Networks. TSTP 2019. Advances in Intelligent Systems and Computing*; Macioszek, E., Sierpiński, G., Eds.; Springer: Cham, Switzerland, 2020; Volume 1083.
19. Dias, G.; Arsenio, E.; Ribeiro, P. The Role of Shared E-Scooter Systems in Urban Sustainability and Resilience during the COVID-19 Mobility Restrictions. *Sustainability* **2021**, *13*, 7084. [CrossRef]
20. Becker, H.; Balac, M.; Ciari, F.; Axhausen, K.W. Assessing the welfare impacts of Shared Mobility and Mobility as a Service (MaaS). *Transp. Res. Part A Policy Pract.* **2020**, *131*, 228–243. [CrossRef]
21. Khatoun, R.; Zeadally, S. Smart cities. *Comm. ACM* **2016**, *59*, 46–57. [CrossRef]
22. Carbonnell, J. Smart-City: Stakeholders Roles and Needs. Available online: https://medium.com/@julien.carbonnell/smart-city-stakeholders-roles-and-needs-8e3679764d2a (accessed on 15 July 2020).
23. Monzón, A. Smart cities concept and challenges: Bases for the assessment of smart city projects. In *Smart Cities, Green Technologies, and Intelligent Transport Systems, Proceedings of the 4th International Conference, SMARTGREENS 2015, and 1st International Conference VEHITS 2015, Lisbon, Portugal, 20–22 May 2015*; Helfert, M., Krempels, K.-H., Klein, C., Donellan, B., Guiskhin, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2015.
24. Martin, B.R.; Irvine, J. *Foresight in Science: Picking the Winners*; Pinter Publishers: London, UK, 1984.
25. Eerola, A.; Jørgensen, B.H. *Technology Foresight in the Nordic Countries*; Rise National Laboratory: Roskilde, Denmark, 2002.
26. Miles, I.; Saritas, O.; Sokolov, A. *Foresight for Science, Technology and Innovation*; Springer: Cham, Switzerland, 2016.
27. Martin, B.R.; Irvine, J. *Research Foresight: Priority-Setting in Science*; Pinter Publishers: London, UK, 1989.
28. Gerdsri, N. Technology foresight concept, activities, challenges and limitations. In *Technology Assessment: Forecasting Future Adoption of Emerging Technologies*; Daim, T.U., Gerdsri, N., Basoglu, N., Eds.; Erich Schmidt: Berlin, Germany, 2010; pp. 131–143.
29. Coates, J.F. Foresight in federal government policy making. *Futur. Res. Q.* **1985**, *1*, 29–53.
30. Coates, J. The needs for new and improved forecasting tools. In Proceedings of the EU–US Scientific Seminar: New Technology Foresight, Forecasting & Assessment Methods, IPTS, Seville, Spain, 13–14 May 2004.
31. Saritas, O.; Oner, M.A. Systemic analysis of UK foresight results joint application of integrated management model and roadmapping. *Technol. Forecast. Soc. Chang.* **2004**, *71*, 27–65. [CrossRef]
32. UK Government Office of Science. *Future of Cities: Foresight for Cities*; UK Government Office of Science: London, UK, 2016.
33. Tewdwr-Jones, M.; Goddard, J. A future for cities? Building new methodologies and systems for urban foresight. *Town Plan. Rev.* **2014**, *85*, 773–794. [CrossRef]
34. Fabbri, E. Strategic planning and foresight: The case of smart specialization strategy in Tuscany. *Foresight* **2016**, *18*, 491–508. [CrossRef]
35. Bina, O.; Ricci, A. Exploring participatory scenario and storyline building for sustainable urban futures—The case of China in 2050. *Foresight* **2016**, *18*, 509–534. [CrossRef]
36. Zahraei, S.M.; Kurniawan, J.H.; Cheah, L. A Foresight study on urban mobility: Singapore in 2040. *Foresight* **2019**, *22*, 37–52. [CrossRef]
37. Güell, J.M.F.; López, J.G. Cities futures. A critical assessment of how future studies are applied to cities. *Foresight* **2016**, *18*, 454–468. [CrossRef]
38. Gerdsri, N. An analytical approach to building a technology development envelope (TDE) for roadmapping of emerging technologies. *Int. J. Innov. Technol. Manag.* **2007**, *4*, 121–135. [CrossRef]
39. Kostoff, R.N.; Schaller, R.R. Science and technology roadmaps. *IEEE Trans. Eng. Manag.* **2001**, *48*, 132–143. [CrossRef]
40. Gerdsri, N.; Puengrusme, S.; Vatananan, R.; Tansurat, P. Conceptual framework to assess the impacts of changes on the status of a roadmap. *J. Eng. Technol. Manag.* **2019**, *52*, 16–31. [CrossRef]
41. Zurcher, R.; Kostoff, R.N. Modeling technology roadmaps. *J. Technol. Transf.* **1997**, *22*, 73–79. [CrossRef]
42. Vatananan, R.S.; Gerdsri, N. The current state of technology roadmapping (TRM) research and practice. *Int. J. Innov. Technol. Manag.* **2012**, *9*, 1250032. [CrossRef]
43. Galvin, R. Science roadmaps. *Science* **1998**, *280*, 803–804. [CrossRef]
44. Phaal, R.; Farrukh, C.; Probert, D. Customizing roadmapping. *Res. Technol. Manag.* **2004**, *47*, 26–37. [CrossRef]
45. Kostoff, R.N.; Boylan, R.; Simons, G.R. Disruptive Technology Roadmaps. *Technol. Forecast. Soc. Chang.* **2004**, *71*, 141–159. [CrossRef]
46. Passey, S.; Goh, N.; Kil, P. Targeting the innovation roadmap event horizon: Product concept visioning & scenario building. In Proceedings of the IEEE International Conference on Management of Innovation and Technology, Singapore, 21–23 June 2006; pp. 604–607.
47. Lizaso, F.; Reger, G. Linking roadmapping and scenarios as an approach for strategic technology planning. *Int. J. Technol. Intell. Plan.* **2004**, *1*, 68–86. [CrossRef]
48. Satangput, P.; Gerdsri, N.; Damrongchai, N. Scenario analysis for identifying the development area of future technologies to combat emerging infectious diseases: APEC efforts. In Proceedings of the 5th International Conference on Management of Innovation and Technology (ICMIT), Singapore, 2–5 June 2010.
49. Lee, S.; Kang, S.; Park, Y.; Park, Y. Technology roadmapping for R&D planning: The case of the Korean parts and materials industry. *Technovation* **2007**, *27*, 433–445.

50. Yasunaga, Y.; Yoon, T. Technology roadmapping with structuring knowledge and its advantages in R&D management. In Proceedings of the IEEE International Engineering Management Conference (IEEE Cat. No. 04ch37574), Singapore, 18–21 October 2004; pp. 581–585.
51. Gerdsri, N.; Kocaoglu, D.F. Applying the analytic hierarchy process (AHP) to build a strategic framework for technology roadmapping. *Math. Comp. Model.* **2007**, *46*, 1071–1080. [CrossRef]
52. Yoon, B.; Phaal, R.; Probert, D. Morphology analysis for technology roadmapping: Application of text mining. *R&D Manag.* **2008**, *38*, 51–68.
53. Chutivongse, N.; Gerdsri, N. Creating an innovative organization: Analytical approach to develop a strategic roadmap guiding organizational development. *J. Model. Manag.* **2020**, *15*, 50–88. [CrossRef]
54. Gerdsri, N.; Manotungvorapun, N. Systemizing the management of university-industry collaboration (UIC): Assessment and roadmapping. *IEEE Trans. Eng. Manag.* **2022**, *69*, 245–261. [CrossRef]
55. Star, J.; Rowland, E.L.; Black, M.E.; Enquist, C.A.F.; Garfin, G.; Hoffman, C.H. Supporting adaptation decisions through scenario planning: Enabling the effective use of multiple methods. *Clim. Risk Manag.* **2016**, *13*, 88–94. [CrossRef]
56. Urbanalyse: Bangkok's Public Transport Network. Available online: https://urbanalyse.com/research/bangkoks-public-network/ (accessed on 3 May 2020).
57. Turo'n, K.; Kubik, A.; Chen, F. When, What and How to Teach about Electric Mobility? An Innovative Teaching Concept for All Stages of Education: Lessons from Poland. *Energies* **2021**, *14*, 6440. [CrossRef]
58. Wangsness, P.B.; Proost, S.; Rødseth, K.L. Optimal policies for electromobility: Joint assessment of transport and electricity distribution costs in Norway. *Util. Policy* **2021**, *72*, 247. [CrossRef]

sustainability

Article

Research on Influential Factors of Satisfaction for Residents in Unit Communities—Taking Ningbo City as an Example

Xing Lian [1,†], Danna Li [1,†], Weifeng Di [1,2], Mohamed Oubibi [1], Xueyan Zhang [3], Sijia Zhang [1], Chengyu Xu [4] and Hejie Lu [1,3,*]

1 College of Teacher Education, College of Education and Human Development, Zhejiang Normal University, Jinhua 321004, China; siro1116@zjnu.edu.cn (X.L.); lidanna@zjnu.edu.cn (D.L.); diwf@zjnu.cn (W.D.); liby478@nenu.edu.cn (M.O.); zeal0224@zjnu.edu.cn (S.Z.)
2 Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, Jinhua 321004, China
3 Lifelong Education Division, Ningbo Open University, Ningbo 315016, China; xyzh@nbou.cn
4 College of Design, Creative and Digital Industries, The University of Westminster, London HA1 3TP, UK; jinghui@zjnu.edu.cn
* Correspondence: luhj@nbou.cn
† These authors contributed equally to this work.

check for updates

Citation: Lian, X.; Li, D.; Di, W.; Oubibi, M.; Zhang, X.; Zhang, S.; Xu, C.; Lu, H. Research on Influential Factors of Satisfaction for Residents in Unit Communities—Taking Ningbo City as an Example. *Sustainability* **2022**, *14*, 6687. https://doi.org/10.3390/su14116687

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 11 April 2022
Accepted: 26 May 2022
Published: 30 May 2022

Abstract: Community governance is the foundation of social governance and a guarantee for modernizing the national governance system and capacity. Residents' satisfaction with community governance is an important factor in measuring community governance level. Based on similar related literature and practical research analysis, 18 influencing factors of community governance residents' satisfaction were extracted from four dimensions. Through the principal component analysis of social science statistical software, 16 key influencing factors were finally selected, and six regression models were analyzed and compared. Further, they clarified the influence of various dimensions on residents' satisfaction with community governance. Further analysis and research on the model show that community governance service, community education and community interpersonal communication are significantly positively correlated with residents' satisfaction with community governance. Although participation in community governance as a single influencing factor has no significant effect in this study, it can also positively and significantly affect residents' satisfaction with community governance under the joint action of community education and community interpersonal communication. Therefore, from the aspects of improving the service level of community governance, enhancing the participation of community governance in multiple ways, giving full attention to the role of community education and shaping the new model of community communication, a more realistic evaluation system of community governance residents' satisfaction is designed.

Keywords: residents' satisfaction with community governance; influential factors; empirical analysis; community governance level

1. Introduction

In the recent Chinese governance system, community governance lays a fundamental position [1]. Community governance is the joint management of community public affairs by the government, community organizations and community citizens [2]. It has become the key point of social governance, the link between government governance and residents' autonomy, which is a fundamental condition to truly promote the modernization of governance in China [3]. The report of the 19th National Congress of the Communist Party of China pointed out that it is necessary to "strengthen the construction of the community governance system and promote the shift of the focus of social governance to the grassroots level". In 2020, General Secretary Xi Jinping pointed out during his inspection at the Jilin Changchun Community Cadre College: "Only by strengthening social governance can we promote the modernization of the national governance system and governance capacity". In April 2021, the Communist Party of China Central Committee and the State Council had released a plan for modernizing the system and capacity for primary-level governance, which shows that we should establish a sound institutional mechanism for grassroots governance, promote healthy interaction between government governance, social regulation and residents' autonomy, and then improve the level of socialization, rule of law, legalization and professionalization of grassroots governance [3].

The level of community governance is directly related to the happiness index of residents and affects social harmony and stability and the effectiveness of social governance to a certain extent. In recent years, China's urban and rural community governance and service level have improved significantly, however, there is still a certain gap in the people's expectations. Community governance must establish the recognition of common goals through negotiation and cooperation and rely on the acceptance and approval of community residents to jointly manage the common affairs of the community, to achieve a high-level and modern community governance system, build a civilized, harmonious and beautiful socialist community with Chinese characteristics. Against this background, this paper first conducts an inductive analysis of the factors affecting residents' community governance satisfaction. Secondly, by using principal component analysis and regression analysis, the key factors affecting residents' satisfaction with community governance are explored from the perspectives of community governance services, resident participation, community education, and community interpersonal communication. At last, according to the research results, this paper provides decision-making and reference for improving community governance capacity from the perspective of practical effects.

This research has positive theoretical and practical significance. Firstly, from the perspective of theoretical significance, this study expands on the theories related to community governance. In view of the community governance, exploring the content and methods of community governance and focusing on the satisfaction of residents in the community can reflect the people-oriented value concept, which is helpful to clarify the role and function of residents in community governance, thus, improves the efficiency of community governance. Secondly, this study enriches the theoretical system of social governance and national governance. Starting from the research theme of "How to improve residents' satisfaction in community governance", this paper sorted out the current state of community governance and discussed the problem of low residents' satisfaction, providing an experience for improving the level of community governance and improving residents' lives. From the perspective of practical significance, the current community governance is in a new stage from single governance to multiple co-governance, which emphasizes the joint governance of the government and the public and pays more attention to the subjective feelings of residents, so as to better improve the work of the community. First of all, this paper focuses on the analysis of residents' satisfaction with community governance in City N and the analysis of existing problems, which can provide a reference for other regions to conduct experimental exploration. Then, it is beneficial to the optimization of resource allocation by the government. By digging out the reasons why current residents are not satisfied with community governance, we can find the shortcomings and weaknesses of government resource allocation and put forward targeted optimization paths, which will encourage the government to further transform its functions and optimize resource allocation.

This study draws lessons from the research results of enterprise satisfaction to study and evaluate residents' satisfaction with community governance. In the relative satisfaction relationship, residents are the objects of community governance services and the beneficiaries of community public services. In this paper, the level of satisfaction is how community residents are satisfied with community governance public services.

2. Core Concepts and Literature Review

2.1. Influencing Factors of Resident Satisfaction with Community Governance

In the study of resident satisfaction, foreign scholars Ladwig and McCallum defined it as "the psychosocial response of residents to community services" [4]. After it, researchers, such as Heller, Rasmussen, Cook and Wolosin, have enriched the concept of resident satisfaction in terms of interpersonal interactions, value systems and life needs. Bard John divided resident satisfaction into general areas and specific areas [5]. Satisfaction in the general domain includes how people experience and feel about the community they live in, while satisfaction in the specific domain includes alienation between people as well as warmth and belonging, how people feel about community administration, and the quality of the environment, individual responsibility, etc. At present, the research on people's satisfaction in China is still in its infancy. There is no clear definition of people's satisfaction and the analyses of various aspects are not perfect enough. However, in terms of practical application, our government attaches great importance to the work of public satisfaction [6].

In the research on residents' satisfaction with community governance, foreign scholars like Rojek believed that community service quality is an important aspect of detecting residents' satisfaction with the community [7]. This service includes education, medical treatment and safety. Meanwhile, Christenson argued that the quality of life of residents in the Chinese community is the main entry point for community satisfaction research [8]. Chinese scholar Jihong put forward that community governance capacity should include community service provision, resident participation, cultural leadership, safety and conflict regulation and community informatization capacity [9]. Community governance capability includes community service supply, resident participation, community cultural leadership, community security and conflict adjustment, and community informatization capability [10]. Therefore, this study combines the actual development of urban and rural communities in China and analyzes resident satisfaction from four aspects: community governance service, community resident participation, community education and community interpersonal communication.

2.1.1. Community Governance Service

Community governance service capability mainly includes the leadership of grassroots party organizations, the public service capabilities of government and community departments, the service capacity of social organizations, and the self-organization capacity of residents [9]. Shiqiang fully affirmed the leading role of the party organization in community autonomy, which can effectively promote the voluntary participation of community residents in community governance [11]. Puqu thought that the structural and functional relationship between community party organizations and neighborhood committees constitutes the basis for their positive interaction, creating conditions for good governance in Chinese urban communities [12]. Feng's views provided new ideas for pluralistic governance and pointed out that under the environment of government purchase policy, the cooperation between social organizations, governments and communities has brought new opportunities for the transformation of community governance. Wei et al. believed that following the three-dimensional logic of "overall guidance of Party organizations and government—co-governance of social organizations and community organizations—the establishment of resident community" can promote the benign interaction of multiple subjects. It can not only enrich the innovative practices of grassroots social governance but also create a good community atmosphere in which people of all ethnic groups live, study, work, and enjoy life together in the new era, and lay a solid ideological foundation for the community of the Chinese nation at the community level [13].

2.1.2. Community Resident Participation

The factors of residents' participation in community governance include residents' willingness and participation degree. The essence of community construction lies in interaction, and residents' participation is at the core of community governance. Jijun believed that the current key work improves residents' participation willingness, guides self-governance, and cultivates community volunteer service spirit [14]. Jing put forward the views on improving the participation of residents in community governance and believed that increasing the participation of residents can effectively improve the level of community governance and enhance the harmony of the community [15]. The countermeasures included strengthening education to improve residents' ability to handle affairs and improving the degree and effectiveness of resident participation [15].

2.1.3. Community Education

The functions of community education include improving the safety awareness of community residents and the organizational degree of the community, mobilizing multiple subjects to participate in community governance, and promoting the formation of a healthy community culture [16]. Dongquan also believed that grassroots party organizations and community education, as the main body and mode of governance, respectively, play important roles in changing the concept of new citizens, improving the quality of residents, building social networks, and strengthening residents' sense of community identity and belonging [17].

2.1.4. Community Interpersonal Communication

Yanfei's research results showed that strengthening community culture and cultivating good interpersonal relationships can help carry out community management work, improve community governance service capabilities, and enhance residents' sense of identity with the community, thereby improving residents' satisfaction with community governance [18]. Xiaoyu thought that positive emotions, such as community emotional interaction and community belonging, could help realize the effective connection between "people" and community, create a community of co-governance, and promote the participation of diverse subjects in community governance [19].

3. Research Design

3.1. Research Tools

Based on the community governance performance questionnaire launched by Hong [20] and combining the actual situation of the research object and the perspective of the research team, this study compiles four scales from four dimensions: community governance services, community resident participation, the impact of community education on community governance, and community interpersonal communication. Specifically, these scales are, respectively, the impact of community governance services on residents' satisfaction with community governance, the impact of residents' participation in community governance on residents' satisfaction with community governance, the impact of community education on residents' satisfaction with community governance, and the impact of community interpersonal communication on residents' satisfaction with community governance. The questionnaire consists of residents' basic information and questions about social governance satisfaction. In addition to the subjective and multiple-choice questions, the Likert five-level scale is used for social governance satisfaction (Appendix A). Participation in social governance is mainly examined through multiple-choice questions, requiring respondents to select the relevant activities they have participated in and score them according to the number of activities they selected (a five-point system). The subjective questions examine the satisfaction of community governance. The study uses python's snownlp package for sentiment analysis, and the scoring results are also divided into five-level scales: the value range between (0.8, 1) represents 5, (0.6, 0.8) represents 4, (0.4, 0.6) represents 3, (0.2, 0.4) represents 2, and (0, 2) represents 1. This study's data analysis and statistical tools are mainly SPSS23.0 and PYTHON.

3.2. Sample Selection

Residents aged 18 and above in more than 20 communities in District B of N City are selected as the research objects, and the survey period is from November 2020 to March 2021. Questionnaires are compiled on an application software called "Wenjuanxing" and distributed to residents through adult schools, community colleges, streets and neighborhood committees. A total of 1056 questionnaires were received. After removing the missing and invalid questionnaires, 1044 were valid, with an effective rate of 98.7%. The basic information of the respondents is as follows (Table 1):

Table 1. Distribution of Survey Objects (n = 1044).

	Category	Number of People	Percentage (%)
Gender	male	342	32.8
	female	702	67.2
Age	0–17	0	0
	18–45	527	50.5
	46–59	314	30.1
	60–74	193	18.5
	>75	10	1.0
Years of Residence	≤1	83	8.0
	(1, 5]	318	30.4
	>5	643	61.6
Political Status	Party member of CPC	371	35.5
	League member	75	7.2
	The Masses	538	51.5
	Democratic parties	39	3.7
	Non	21	2.0
Education	Junior high school and below	297	28.4
	Senior high school (secondary vocational school)	235	22.5
	Junior college	239	22.9
	Undergraduate	261	25.0
	Graduate and above	12	1.1
Occupation	Heads of state organs, party and mass organizations, enterprises and institutions	123	11.8
	Professional technicians	194	18.6
	Office staff and related personnel	243	23.3
	Business and service personnel	241	23.1
	Production personnel of agriculture, forestry, animal husbandry, fishery and water conservancy	71	6.8
	Soldier	4	0.4
	Other	168	16.1
Monthly Income	3000 and below	199	19.1
	3000–5000	421	40.3
	5000–8000	274	26.2
	8000–15,000	124	11.9
	15,000–30,000	24	2.3
	More than 30,000	2	0.2

3.3. Research Approach

SPSS23.0 is used in this study to conduct a statistical analysis on the collected questionnaire data. Firstly, we optimized the research tools and improved the original questionnaire by issuing test questionnaires. The Likert-style questions were changed to subjective questions, and individual engagement questions were changed to multiple-choice questions. Excluding individual invalid and highly interfering questions (Q15, Q16), the overall Cronbach α coefficient is 0.958, indicating that the corresponding variables of each factor have strong internal consistency and the scale has good reliability.

Secondly, SPSS was used to convert all factors into standardized values, and principal component analysis was used to analyze each first-level dimension and determine all factors of each dimension.

Thirdly, we scored each first-level dimension, analyzed the research results, and evaluated the impact of various factors of social governance on the satisfaction of residents in community governance to give corresponding suggestions.

4. Research Results

4.1. Descriptive Statistics

Resident satisfaction with community governance is mainly influenced by community governance services, resident participation, community education and community interpersonal communication (as shown in Table 2).

Table 2. Influencing Factors of Each Dimension (n = 1044).

Influencing Factor Dimension	Influence Factor	Average Value	Standard Deviation
Community Governance Services	Q3 Ask the neighborhood committee if there is a problem	3.86	0.972
	Q4 Strong property management serviceability and good service attitude	3.81	0.977
	Q5 Community party organizations are enthusiastic about serving residents	4.06	0.932
Resident Participation	Q1 Often participate in community activities	3.14	0.852
	Q2 What activities have you participated in	1.82	1.134
	Q12 Willing to participate in community governance	4.11	0.894
	Q13 Willing to participate in community affairs and problem-solving	4.08	0.905
The Role of Community Education	Q9 Community education promotes neighborhood relations	4	0.948
	Q10 Community education promotes community culture construction	4.1	0.914
	Q11 Community education improves the ability of community governance participation	4.13	0.91
Community Interpersonal Communication	Q8 Good neighborhood communication	3.93	0.968
	Q14 Neighbors help each other well	3.802	1.112
Resident Satisfaction of Community Governance	Q6 Public security and order	4.07	0.95
	Q7 Sanitary environment and landscape greening	3.7	1.075
	Q17 Community affairs disclosure and publicity	2.02	0.963
	Q18 Overall satisfaction	3.01	0.234

4.2. Variable Analysis

4.2.1. Dependent Variable: Resident Satisfaction with Community Governance

Resident satisfaction with social governance services is measured through four items (dependent variable S in Table 3): Questions 6, 7, 17 and 18. Among them, Question 18 is a subjective question, which is scored using the SnownLP package of Python, and the score after processing is 1–5 points. The answers to the remaining questions are assigned 1–5 points from "strongly disagree" to "strongly agree". Principal component analysis is used to extract two common factors with eigenvalues greater than 1: "residents' satisfaction with community environmental governance" and "residents' satisfaction with transaction processing and publicity" (KMO = 0.648, $p < 0.001$). Among them, the variance contribution rate of "residents' satisfaction with community environmental governance" is 43.306%, and the variance contribution rate of "residents' satisfaction with transaction processing and publicity" is 25.754%. The cumulative variance contribution rate of the two is 70.060%.

Table 3. Variable Description of Each Dimension (n = 1044).

Dependent Variable						
Variable	**Component Matrix**					
Resident satisfaction with community governance (S) Y = 0.43306 × Y1 + 0.25754 × Y2 KMO = 0.648, $p < 0.001$ Cumulative Variance Contribution Rate = 70.060%	Principal Component 1 (Residents' satisfaction with community environmental governance)					
	Item	Component	Item	Component		
	Q6	0.921	Q17	−0.154		
	Q7	0.922	Q18	0.101		
	Eigenvalue = 1.732, Percentage Variance = 43.306% Y1 = 0.70012 × Q6 + 0.70017 × Q7 − 0.1169 × Q17 + 0.07698 × Q18					
	Principal Component 2 (Residents' satisfaction with transaction handling and publicity of community affairs)					
	Item	Component	Item	Component		
	Q6	0.064	Q17	0.712		
	Q7	−0.027	Q18	0.748		
	Eigenvalue = 1.070, Percentage Variance = 25.754% Y2 = 0.06142 × Q6 − 0.026Q7 + 0.68802 × Q17 + 0.72262 × Q18					
Independent Variable						
Community Governance Services (G) X = X1 KMO = 0.748, $p < 0.001$ Cumulative Variance Contribution Rate = 81.814%	Principal Component 1 (Community governance services)					
	Item	Component	Item	Component	Item	Component
	Q3	0.909	Q4	0.902	Q5	0.902
	Eigenvalue = 2.454, Percentage Variance = 81.814% X1 = 0.58046 × Q3 + 0.57604 × Q4 + 0.57604 × Q5					
Resident participation in community governance (P) X = 0.51663 × X1 + 0.30820 × X2 KMO = 0.630, $p < 0.001$ Cumulative Variance Contribution Rate = 82.484%	Principal Component 1 (Willingness to participate in community governance)					
	Item	Component	Item	Component		
	Q1	0.461	Q12	0.912		
	Q2	0.448	Q13	0.907		
	Eigenvalue = 2.067, Percentage Variance = 51.663% X1 = 0.32055 × Q1+ 0.31172 × Q2+ 0.63436 × Q12 + 0.63060 × Q13					
	Principal Component 2 (Community governance participation)					
	Item	Component	Item	Component		
	Q1	0.697	Q12	−0.344		
	Q2	0.708	Q13	−0.358		
	Eigenvalue = 1.233, Percentage Variance = 30.820% X2 = 0.79213 × Q1 + 0.79841 × Q2 − 0.55651 × Q12 − 0.56791 × Q13					

Table 3. *Cont.*

<table>
<tr><th colspan="7">Dependent Variable</th></tr>
<tr><th>Variable</th><th colspan="6">Component Matrix</th></tr>
<tr><td rowspan="4">Effect of community education on community governance (E)
X = X1
KMO = 0.755, $p < 0.001$
Cumulative Variance Contribution Rate = 88.852%</td><td colspan="6">Principal Component 1 (The role of community education in community governance)</td></tr>
<tr><td>Item</td><td>Component</td><td>Item</td><td>Component</td><td>Item</td><td>Component</td></tr>
<tr><td>Q9</td><td>0.933</td><td>Q10</td><td>0.958</td><td>Q11</td><td>0.938</td></tr>
<tr><td colspan="6">Eigenvalue = 2.666, Percentage Variance = 88.852%
X1 = 0.57116 × Q9 + 0.58647 × Q10 + 0.57430 × Q11</td></tr>
<tr><td rowspan="4">Community Interpersonal Communication (C)
X = X1
KMO = 0.612, $p < 0.001$
Cumulative Variance Contribution Rate = 88.018%</td><td colspan="6">Principal Component 1 (Community interpersonal communication)</td></tr>
<tr><td>Item</td><td colspan="2">Component</td><td>Item</td><td colspan="2">Component</td></tr>
<tr><td>Q8</td><td colspan="2">0.938</td><td>Q14</td><td colspan="2">0.938</td></tr>
<tr><td colspan="6">Eigenvalue = 1.760, Percentage Variance = 88.018%
X1 = 0.70711 × Q8 + 0.70711 × Q14</td></tr>
<tr><td colspan="7">Control Variables</td></tr>
<tr><td>Variables</td><td colspan="6">Explain</td></tr>
<tr><td>Gender</td><td colspan="6">Male = 1; Female = 2</td></tr>
<tr><td>Age</td><td colspan="6">The age range from low to high is 1–5</td></tr>
<tr><td>Years of Residence</td><td colspan="6"><1 year is 1, 1–5 years is 2, and more than 5 years is 3</td></tr>
<tr><td>Political Status</td><td colspan="6">1 = “Party Member of CPC”; 2 = “League Member”; 3 = “The Masses”; 4 = “Democratic Parties”; 5 = “Non”;</td></tr>
<tr><td>Educational Background</td><td colspan="6">1 = “Junior high school and below”; 2 = “Senior high school (Secondary vocational)”; 3 = “Junior college”; 4 = “Undergraduate”; 5 = “Graduate and above”</td></tr>
<tr><td>Occupation</td><td colspan="6">1 = “Person in charge of state organs, Party mass organizations, Enterprises and institutions”; 2 = “Professional technicians”; 3 = “Office staff and related personnel”; 4 = “Business and service personnel”; 5 = “Production personnel in agriculture, Forestry, Animal husbandry, Fishery and water conservancy”; 6 = “Soldier”; 7 = “Others”</td></tr>
<tr><td>Monthly Income</td><td colspan="6">1 = “3000 and below”; 2 = “3000–5000”; 3 = “5000–8000”; 4 = “8000–15,000”; 5 = “15,000–30,000”</td></tr>
</table>

4.2.2. Independent Variables

Independent variable I: Community governance service (independent variable G in Table 3). Mainly including Questions 3–5. The principal component analysis method extracts a common factor, “community governance service”, with an eigenvalue of 2.454 (KMO = 0.748, $p < 0.001$), and its variance contribution rate is 81.814%.

Independent variable II: Resident participation in community governance (independent variable P in Table 3). Questions 1, 2, 12 and 13 are included. The principal component analysis is used to extract two common factors with eigenvalues greater than 1: “Willingness to participate in community governance” (Table 3, independent variable P, principal component 1) and “participation in community governance” (Table 3, independent variable P, principal component 2). The cumulative variance contribution rate of the two is 82.484%. According to the variance contribution rate, it is synthesized into “resident participation in community governance” (KMO = 0.630, $p < 0.001$).

Independent variable III: Effect of community education on community governance (Table 3, independent variable E). Questions 9, 10 and 11 are involved. The principal component, “community education promotes community governance” (KMO = 0.755, $p < 0.001$), is extracted by principal component analysis, and its variance contribution rate is 88.852%.

Independent variable IV: Community interpersonal communication (Table 3, independent variable C). Question 8 and Question 14 are designed to test this factor. One principal component, “community interpersonal communication”, is extracted by principal component analysis (KMO = 0.612, $p < 0.001$), and its variance contribution rate is 88.018%.

4.2.3. Control Variables

In this study, gender, age, years of residence, political status, educational background, occupation and monthly income are taken as control variables to investigate their impact on residents' satisfaction with community governance. All variables are numbered nominal variables, and the corresponding relationship between the specific attributes and values of variables can be referred to in Tables 2 and 3.

4.3. Regression Analysis Model of Influencing Factors of Residents' Satisfaction in Community Governance

In this paper, resident satisfaction with community governance is used as a dependent variable, and community governance service, resident participation in community governance, the effect of community education on community governance, and community interpersonal communication are used as independent variables for regression analysis. The variance inflation factor VIF of each explanatory variable is less than 4.6, so there is no multicollinearity among the variables in the model, and its Durbin–Watson coefficient is 1.925, indicating that there is no autocorrelation in the residuals [21]. Model 1 mainly observes the impact of individual-level factors on residents' satisfaction with community governance. Models 2, 3, 4, 5, and 6 examine the impact of each variable on the satisfaction of residents with community governance. Model 6 includes all independent variables and control variables.

4.3.1. The Influence of Individual Factors on Resident Satisfaction with Community Governance

From Model 1, it can be seen that the years of residence and political status have a significant negative impact on residents' satisfaction with community governance, while other individual-level factors have no significant impact on it. Further investigation and data analysis found that residents with longer living years mostly lived in old communities, and their living environment and governance concepts were relatively weak. As a result, the coefficient was −0.003 *. Party members have relatively high levels of satisfaction with community governance and found that grassroots party organizations have played a great role in it, while those without party affiliation and democratic parties are relatively low, and it is found that these people have a relatively high degree of education. The research of relevant scholars also indicates that residents with high educational backgrounds have higher requirements and expectations for community governance [22].

4.3.2. The Influence of Each Independent Variable on Residents' Satisfaction with Community Governance

Model 2 adds community governance service factors based on Model 1. The results show that community governance services are significantly positively correlated with residents' satisfaction with community governance (0.317 ***), which can explain about 66% of the variation in satisfaction, indicating that the better the community governance service, the higher the residents' satisfaction with the community governance.

Model 3 adds community governance participation factors based on Model 1. The results show that residents' satisfaction with community governance is significantly positively correlated (0.366 ***), but it can only explain 23.8% of the variation in satisfaction, so community governance participation is not statistically significant here.

Model 4 adds the role of community education based on Model 1. The results show that the integration of community education into community governance has a significant positive correlation (0.319 ***) with residents' satisfaction with community governance, and it can explain about 68% of the variation in satisfaction. Therefore, the integration of community education into community governance positively affects residents' satisfaction with community governance's greater significance.

Model 5 adds community interpersonal factors to Model 1. The results show that community interpersonal communication factors are significantly positively correlated with resident satisfaction with community governance (0.322 ***), and it can explain nearly 59%

of the variation in satisfaction. Therefore, the more harmonious community interpersonal relationships are, the more conducive to improving resident satisfaction.

Model 6 is the full model. The data in Table 4 show that the factors of community governance service, resident participation in community governance, the effect of community education on community governance, and community interpersonal communication are all significantly positively correlated with resident satisfaction of community governance, and they can explain 75% of the variation in satisfaction. Among them, the coefficients of community governance service, community education and community interpersonal communication all exceed 0.1, indicating that these three factors have a greater impact on resident satisfaction in community governance, while community governance participation is relatively weak.

Table 4. Analysis of the Influence Factors of Various Dimensions on Resident Satisfaction with Community Governance.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Years of Residence	−0.003 *	−0.004 ***	−0.003 *	−0.004 ***	−0.003 **	−0.004 ***
Gender	−0.037	0.030	−0.058	−0.042	0.012	0.003
Age	−0.047	0.026	−0.059 **	0.010	0.019	0.026 *
Political Status	−0.048 *	−0.002	0.030	0.006	−0.020	0.009
Educational Background	−0.012	0.004	−0.019	0.013	0.009	0.011
Occupation	−0.009	0.006	−0.001	−0.001	0.004	0.005
Monthly Income	0.029	0.004	0.069 ***	0.009	0.023	0.013
Community Governance Services		0.317 ***				0.133 ***
Resident Participation			0.366 ***			0.034 *
The Role of Community Education				0.319 ***		0.124 ***
Community Interpersonal Communication					0.322 ***	0.105 ***

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Model 1: $R^2 = 0.026$, $\triangle R^2 = 0.020$, F = 3.993, sig < 0.001; Model 2: $R^2 = 0.664$, $\triangle R^2 = 0.661$, F = 255.174, sig < 0.001; Model 3: $R^2 = 0.243$, $\triangle R^2 = 0.238$, F = 41.626, sig < 0.001; Model 4: $R^2 = 0.683$, $\triangle R^2 = 0.681$, F = 279.257, sig < 0.001; Model 5: $R^2 = 0.588$, $\triangle R^2 = 0.585$, F = 184.562, sig < 0.001; Model 6: $R^2 = 0.753$, $\triangle R^2 = 0.751$, F = 286.255, sig < 0.001.

5. Conclusions and Countermeasures

5.1. Conclusions

This paper studies residents' satisfaction with community governance from four dimensions: community governance service, community resident participation, community education and community interpersonal communication. Firstly, through the literature review and in-depth interviews with community staff and residents, teachers in adult schools and community colleges, 18 influencing factors are extracted. Secondly, SPSS is used to conduct a principal component analysis of the questionnaire, which obtained the four dimensions mentioned above and 16 influencing factors. Finally, regression analysis analyzes the relationship between explanatory and dependent variables. The main conclusions are as follows.

Firstly, community governance service, community education, and community interpersonal communication significantly correlate with residents' satisfaction with community governance. During the individual interviews, some residents mentioned that "the community now has easy access to medical services, sports venues and educational venues, so they are very satisfied". High-level community governance services can help improve residents' satisfaction with community governance. Community education plays a greater role in improving residents' literacy and participation in community governance and publicizing community governance, thereby promoting resident satisfaction. Harmonious and mutual helping community interpersonal communication can improve residents' happiness, affect the implementation of specific policies and measures of community governance, and positively affect residents' satisfaction with community governance.

Secondly, as an independent influencing factor, community governance participation has no significant effect in this study; however, it does not mean that community governance participation is meaningless to resident satisfaction. Under the joint influence of variables, such as community education and community interpersonal communication, community governance participation can significantly affect residents' satisfaction with community governance. The empirical research of Gao Hong et al. showed that community governance participation has the most significant positive impact on community governance performance among all the influencing factors.

5.2. Countermeasures

Based on the above conclusions, the following countermeasures are proposed:

First and foremost, improve the level of community governance services. On the one hand, establish a sound, scientific and reasonable organizational structure for community governance, including community party organization, community development coordination committee, and a resident's supervision committee. The community party branch should play the role of "axis", and vertically promote the "community party committee-community party branch-building party group" structure by coordinating the implementation and strengthening of the internal integration capacity of the party organization. Horizontally, the cooperative governance mechanism needs to be improved to mobilize party organizations, community organizations and profit-making organizations to jointly participate in grassroots governance services [23]. On the other hand, they should give full attention to the role of community interpersonal communication and use the grid governance system to promote the implementation of community governance [24]. According to the logic of "one core, one grid, multi-governance and co-operative scenario", primary, secondary and tertiary grids are divided to further enhance the effectiveness of grassroots governance by taking the grid as the coverage area and the community residents as the central point. Finally, explore integrating party construction work and community governance to improve community governance services so as to further improve the system of grassroots mass self-governance led by grassroots party organizations. For issues of concern to the community and major community affairs, the party organization should take the lead in holding listening and evaluation sessions.

Secondly, mobilize the enthusiasm of residents to participate in community governance. First, establish a mechanism and channel for community party and league members to participate in community governance and service, actively gather the backbone of the community, and make them exert their functions. By setting up party groups and building leaders in a piecemeal manner, "horizontally to the edge and vertically to the bottom" [25], the tentacles of party-residence organizations are further extended to allow party members and group members to play a role in community governance [26]. Secondly, make full use of the advantages of social organizations and establish an effective path to improve the level of community governance [27]. As a vehicle for residents to participate in community activities, community social organizations promote neighborhood deconfliction and the creation of new customs based on the promotion of collaboration in knowledge and skills among the organization's members [28]. At the same time, it is necessary to give full access to the positive role of community education in establishing and maintaining social organizations to improve community governance ability and personal literacy of community residents. On the one hand, the integration of various educational resources in the community promotes the organic integration of all levels and types of education; on the other hand, it enhances the sense of participation in residents' subjects, fosters learning organizations and groups, and builds public learning spaces, so as to give better access to the role of community education work in promoting the process of community governance.

Thirdly, give full access to the function of community education. The first step is to explore participation channels to establish and improve the system and mechanism of community education participation in community governance. The community education co-management model of "party-building-led, government-led and multi-party participation" should be adhered to [29]. A three-tier network of "district-level community colleges-street branches-community teaching points" was supposed to be further established in order to better expand the coverage of community teaching points in the community. The second step is to optimize and expand the contents and forms of community education. In terms of educational content, civic education, education on the rule of law, science education and other learning content that is popular with community residents, can be added. Meanwhile, a scientific evaluation system should be formulated to promote the modernization of community governance and the effectiveness of residents' participation in community education evaluation. Finally, give access to the leading role of community education. It is necessary to give full access to the resource advantages of community education, provide systematic and characteristic services, guide community residents to pay attention to the key and difficult issues of community governance, mobilize multiple subjects to participate in community governance, to pool their wisdom and strength, improve the level of community governance, and enhance resident satisfaction.

Fourthly, shape a new model of community interaction using the internet as a platform. Given the current weak interpersonal communication among community residents, creating a brand new form of interpersonal communication in small and micro diverse communities in the internet era is significant to form harmonious, friendly and mutually helpful community interpersonal relationships [30]. Then, define neighborhood relations as the basis. In view of the fact that many communities—especially old communities—are "unmanaged", "unmanageable" and "unpaid", the current situation is to adopt a "voluntary service + organizational linkage + good neighbor activities" operational mechanism, actively play the role of party members as pioneers, and encourage community residents to participate in the management of daily affairs. This is to achieve the goal of "self-governance to mobilize the masses participating in community affairs, the rule of law to resolve conflicts" in order to smooth the service of "the last meter" for the community masses.

This study also has the following limitations. Firstly, the sample was selected from a more restricted range. Due to the epidemic, the author was unable to reach the communities in other areas to conduct the survey. Therefore, the data collected in this paper has spatial and temporal limitations. Secondly, this paper did not adequately discuss objective environmental differences in communities. As resident satisfaction with community governance may be influenced by the facilities and environmental conditions of the community, future research could collect new data from a large sample of people across the country and test the correlation and causality between the factors that have been generalized by controlling for environmental variables. This could be achieved from a quantitative research perspective, with a view to constructing a more generalized model of resident satisfaction with community governance.

Author Contributions: Conceptualization, X.L., M.O., S.Z., C.X. and H.L.; data curation, D.L., M.O., S.Z., C.X. and H.L.; formal analysis, D.L., X.L., M.O., X.Z. and C.X.; investigation, D.L., X.Z. and X.L.; methodology, M.O. and H.L.; project administration, W.D. and H.L.; resources, W.D.; software, M.O.; supervision, W.D.; validation, W.D.; visualization, H.L.; writing—original draft, H.L.; writing—review and editing, D.L., X.L., M.O., S.Z., X.Z. and C.X. All authors have read and agreed to the published version of the manuscript.

Funding: Zhejiang Province Education Science Planning 2021 Annual Planning Project (Colleges and Universities), "Practice Research on Integrating Community Education in Communities Based on Learning Community" Project number "2021SCG095"; Ningbo Social Science Research Base Project "Cultivating Community Learning Community Learning Community to Lead the Practice Research in Grass-Roots Communities Governance", Project number "2021JD5-PY45".

Institutional Review Board Statement: The study was conducted in accordance with the declaration of Helsinki, and approved by the Ethics Committee of the local university, Ningbo Open University (Protocol code: ZSRT2022018) approved in 1 February 2022.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study and written informed consent has been obtained from the subjects to publish this paper.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Part I: Basic Information				
Which community do you live in?				
A. Lily Comminity **B.** Dagang Community **C.** Cuckoo Community **D.** Lotus Community **E.** Gaotang Community **F.** Begonia Community **G.** Hongmei Community **H.** Lingxiao Community **I.** Rose Community **J.** Milan Community **K.** Sunny Community **L.** Tong Wan Community **M.** Sunny Community **N.** Snow lotus Community **O.** Silver Community **P.** Winter jasmine Community **Q.** Magnolia Community **R.** Zhilan Community **T.** Bauhinia Community **U.** Others				
For how many years you have lived in this community?				
_____Years				
What's your gender?				
A. Male **B.** Female				
What's your age?				
A. 0–14	**B.** +14–45	**C.** 46–59	**D.** 60–74	**E.** 75 and above
Political status?				
A. Party member	**B.** Democratic	**C.** The masses	**D.** League members	**E.** Non-party democrats
Educational background?				
A. Junior and below	**B.** High School	**C.** College	**D.** Undergraduate	**E.** Ms and above
What's your job?				
A. Technical job	**B.** Business, Services	**C.** Soldier	**D.** Production personnel	**E.** Others
Monthly income				
A. less than 3000	**B.** 3 k–5 k	**C.** 8 k–15 k	**D.** 15 k–30 k	**E.** +30.000
Part II: community governance satisfaction problem				
How often do you participate in community activities?				
A. Regularly participate	**B.** Occasionally participate	**C.** Neutral	**D.** seldom participate	**E.** Never participated
What types of community activities have you participated in before?				
A. Lectures on health knowledge	**B.** Folk customs, festivals and other activities	**C.** Community school classes	**D.** Volunteer services	**E.** Others
Do you understand the work of the neighborhood Committee?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree

Do you report problems to the community?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
When you encounter problems in the community, the first thing you will think of through the street neighborhood committee property police station to solve?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
Do you think your community is very safe?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
Do you think the hardware environment of the community is very good?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
Do you think your community is very harmonious and the communication is very good?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
Do you think the community education promotes neighborhood relations?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
Do you think it would be better to participate in community schools and publicize the installation of garbage sorting elevators and epidemic prevention through community schools?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
Do you prefer participating in community activities to improving myself by attending community schools?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
You are very willing to participate in community governance?				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
You are more willing to participate in community problem solving or negotiation.				
A. Strongly Agree	**B.** Agree	**C.** Neutral	**D.** Disagree	**E.** Strongly disagree
What forms of community governance activities have you or your family members (roommates) participated in?				
A. Neighborhood committee election	**B.** Community building activities	**C.** Community volunteers	**D.** Others	**E.** Never participated
What do you think are the reasons why residents don't want to participate in community governance?				
A. Residents have a weak sense of belonging to the community	**B.** Residents percieve themselves to be inadequate	**C.** There are no Channels and platforms for residents to participate.	**D.** They have no time or energy	**E.** Other reasons

In your opinion, which aspects can improve the enthusiasm of citizens to participate in community governance?				
A. Raise awareness of the importance of participation among residents	**B.** Open and transpaent financial and election procedures	**C.** Timely release of important information to residents	**D.** Give residents more space to participate	**E.** Others
You can get the basic street information by? Type of social media				
A. Wechat official account	**B.** Community area manager	**C.** Community activities space	**D.** Community leaders	**E.** Others

Note: Please refer to the corresponding authors for the original version of the questionnaire in Chinese.

References

1. Chen, Y. Grassroots social governance innovation and community economic development pathways. *J. Shanxi Univ. Financ. Econ.* **2022**, *44*, 11–13.
2. Wei, N. China's urban community governance model: Developmental evolution and institutional innovation. *J. Renmin Univ. China* **2003**, *1*, 135–140.
3. Yu, H.Y. Research on the Influence Mechanism of Urban Community Public Service Satisfaction on Residents' Happiness. Ph.D. Thesis, Huazhong University of Science and Technology, Wuhan, China, 2019.
4. McCalm, L. Tackling Conflicts of Interest: Policy Instruments in Different Settings. *Public Integr.* **2016**, *18*, 83–100.
5. Rasmussen, H. Optimizing identification and management of COPD patients—Reviewing the role of the community pharmacist. *Br. J. Clin. Pharmacol.* **2014**, *83*, 192–201.
6. Fute, A.; Oubibi, M.; Sun, B.; Zhou, Y.; Xiao, W. Work Values Predict Job Satisfaction among Chinese Teachers during COVID-19: The Mediation Role of Work Engagement. *Sustainability* **2022**, *14*, 1353. [CrossRef]
7. Rojek, D.G.; Clemente, F.; Summers, G.F. Community Satisfaction: A Study of Content with Local Serbices. *Rural. Sociol.* **1975**, *40*, 177–200.
8. Christenson, J.A. Urbanism and Community Sentiment: Extending Wirth's Model. *Soc. Sci. Q.* **1979**, *60*, 387–400.
9. Ye, J.H. Governance ability of rural-to-residential communities: Dimensions, influencing factors and improvement paths. *Zhongzhou Acad. J.* **2021**, *2*, 59–65.
10. State Council. Opinions on Strengthening and Improving Urban and Rural Community Governance. Available online: http://www.gov.cn/zhengce/2017-06/12/content_5201910.htm (accessed on 12 June 2017).
11. Wang, S.Q. Building a community: The Effective Path of Promoting Party Building and Leading Community Autonomy in the New Era. *Strives Realism* **2021**, *42*, 110–152.
12. Wang, P.J.; Li, T.L. Research on the Interrelationship between Party Organizations and Neighborhood Committees in Chinese Urban Communities—Based on Community Case Analysis in C, N and B Cities. *Hebei J.* **2021**, *41*, 191–199.
13. Li, W.; Li, Z.Y. The internal mechanism and realization path of ethnic inter-embedded communities from the perspective of social governance community. *J. Northwest Univ. Natl. Philos. Soc. Sci. Ed.* **2021**, 80–87. [CrossRef]
14. Zhang, J.J. Guide social forces to participate in community governance. *Journal of Social Sciences*, 22 July 2021.
15. Bai, J. Research on the Countermeasures of Improving Residents' Participation in the Renovation of Old Communities. Master's Thesis, Harbin Normal University, Harbin, China, 2021.
16. Wang, W. Strengthening Community Education to Promote the Modernization of Community Governance in the New Era. *China Adm.* **2021**, *3*, 155–156.
17. Li, D.Q. Conditions, Mechanisms and Paths for Community Governance to Promote the Realization of Citizenization in the Process of New Urbanization. *J. Tongji Univ. Soc. Sci. Ed.* **2021**, *32*, 82–91.
18. Wang, Y.F. Group Work Intervention of Community Identity for the Elderly in Village Transfer. Ph.D. Thesis, Yangzhou University, Yangzhou, China, 2021.
19. Li, X.Y. Community Emotions: Construction of a New Rural Community Governance Community. Ph.D. Thesis, Tianjin University of Technology, Tianjin, China, 2021.
20. Gao, H.; Sun, B.C. The measurement dimension of community publicity and community governance performance—An empirical analysis based on Q city. *Dongyue Lun Cong* **2021**, *42*, 104–113, 191–192.
21. Wang, G.C.; Akabay, C.K. Autocorrelation: Problems and solutions in regression modeli. *J. Bus. Forecast.* **1994**, *13*, 18.
22. Lu, J.J. Neighborhood Space, Institutional Capacity and Community Governance Performance in the Context of Social Transformation. Ph.D. Thesis, Fudan University, Shanghai, China, 2013.
23. Kong, F.Y.; Ruan, H.W. Mobilisation, embedding and integration: Three mechanisms for party organisations to lead grassroots social governance. *Learn. Pract.* **2022**, *2*, 83–91.

24. Huang, W.D. Improving the Level of Community Governance from the Three-Dimensional Level of Architecture, Platform and Mechanism. *People's Forum*, 9 May 2017. Available online: http://www.rmlt.com.cn/2017/0509/473568.shtml (accessed on 10 April 2022).
25. Hu, X.J. Enhance community governance at the three-dimensional level of structure, platform and mechanism. *People's Forum* **2017**, *42*. [CrossRef]
26. Li, X.Z. New exploration of in-service party members participating in community governance. *Frontline* **2020**, 83–85. [CrossRef]
27. Li, Y.X. Content differentiation and organizational optimization of urban residents' community participation: A comparative analysis based on the case of B community and Q community in L city. *Urban Issues* **2020**, 84–90. [CrossRef]
28. Yuan, F.C. Empower residents: The logic of subjectivity and the path of action for community participation. *Exec. Forum* **2019**, *26*, 80–85.
29. Lou, M.Q. Research on community education management in the context of social governance. *Educ. Rev.* **2015**, *9*, 50–52.
30. Tang, X.Y.; Zhang, J.D. "Micro-governance" of urban communities and the transformation of interpersonal interaction patterns in communities. *Soc. Sci.* **2018**, 79–90. [CrossRef]

sustainability

Article

CO_2 Emission Reduction Potential of Road Transport to Achieve Carbon Neutrality in China

Jieshuang Dong, Yiming Li, Wenxiang Li * and Songze Liu

Business School, University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai 200093, China; dongjs@usst.edu.cn (J.D.); 202531033@st.usst.edu.cn (Y.L.); 1935021716@st.usst.edu.cn (S.L.)
* Correspondence: liwx@usst.edu.cn

Abstract: Under the targets of peaking CO_2 emissions and carbon neutrality in China, it is a matter of urgency to reduce the CO_2 emissions of road transport. To explore the CO_2 emission reduction potential of road transport, this study proposes eight policy scenarios: the business-as-usual (BAU), clean electricity (CE), fuel economy improvement (FEI), shared autonomous vehicles (SAV), CO_2 emission trading (CET) (with low, medium, and high carbon prices), and comprehensive (CS) scenarios. The road transport CO_2 emissions from 2020 to 2060 in these scenarios are calculated based on the bottom-up method and are evaluated in the Low Emissions Analysis Platform (LEAP). The Log-Mean Divisia Index (LMDI) method is employed to analyze the contribution of each factor to road transport CO_2 emission reduction in each scenario. The results show that CO_2 emissions of road transport will peak at 1419.5 million tonnes in 2033 under the BAU scenario. In contrast, the peaks of road transport CO_2 emissions in the CE, SAV, FEI, CET-LCP, CET-MCP, CET-HCP, and CS scenarios are decreasing and occur progressively earlier. Under the CS scenario with the greatest CO_2 emission reduction potential, CO_2 emissions of road transport will peak at 1200.37 million tonnes in 2023 and decrease to 217.73 million tonnes by 2060. Fuel structure and fuel economy contribute most to the emission reduction in all scenarios. This study provides possible pathways toward low-carbon road transport for the goal of carbon neutrality in China.

Keywords: road transport CO_2 emissions; Low Emissions Analysis Platform (LEAP); Log-Mean Divisia Index (LMDI) method; carbon neutrality

Citation: Dong, J.; Li, Y.; Li, W.; Liu, S. CO_2 Emission Reduction Potential of Road Transport to Achieve Carbon Neutrality in China. *Sustainability* **2022**, *14*, 5454. https://doi.org/10.3390/su14095454

Academic Editor: Marco Raugei

Received: 20 March 2022
Accepted: 28 April 2022
Published: 1 May 2022

1. Introduction

With improvements in people's living standards, China's vehicle population is increasing yearly. Consequently, CO_2 emissions from road transport are showing an increasing trend. According to data from the International Energy Agency, in 2019, CO_2 emissions generated by the global transport sector accounted for approximately 1/4 (24.2%) of total global CO_2 emissions, with emissions of approximately 8 million tonnes, of which emissions in the road transportation sector were approximately 6.5 billion tonnes, accounting for 81% of the whole transportation sector [1]. In addition, CO_2 emissions from the transport sector have increased at the fastest rate in recent decades. This proportion is expected to increase to 41% by 2030. To address global climate change, China proposed to achieve the goals of CO_2 emission peaking by 2030 and carbon neutrality by 2060. Therefore, reducing transport CO_2 emissions is essential for achieving the goal of carbon neutrality [2,3], especially for the road transport sector.

Due to the uncertainty in policies, the future trend of road transport CO_2 emissions is ambiguous. Therefore, it is a matter of urgency to make reliable predictions on road transport CO_2 emissions and propose effective ways toward low-carbon road transport. Presently, most existing studies focus on the CO_2 emissions of transportation companies, disregarding private transportation. In addition, certain emerging scenarios, such as electric vehicles with cleaner power, shared autonomous vehicles, and CO_2 emission trading markets, are rarely considered. Therefore, research on road transport CO_2 emissions under the target of carbon neutrality is still insufficient.

To fill these gaps, this study aims to explore the CO_2 emission reduction potential of road transport to achieve carbon neutrality in China. First, we propose eight policy scenarios: the business-as-usual (BAU), clean electricity (CE), fuel economy improvement (FEI), shared autonomous vehicles (SAV), CO_2 emission trading (CET) (with low, medium, and high carbon prices), and comprehensive (CS) scenarios. Second, we analyze the CO_2 emission reduction of road transport in each scenario and identify the key factors that contribute to emission reduction. Last, we make several policy suggestions to reduce CO_2 emissions of road transport for the goal of carbon neutrality in China. The main contributions of this study are as follows:

- First, taking into account private road transport, the total CO_2 emissions of China's road transport can be calculated and predicted more comprehensively and accurately.
- Second, future policy scenarios with emerging technologies and markets that are aimed at significantly enhancing the CO_2 emission reduction potential of road transport are introduced.
- Third, the contribution of key factors that influence road transport CO_2 in different scenarios are decomposed to support policy design and decision-making.

The remainder of this study is organized as follows: Section 2 reviews existing research on road transport CO_2 emissions. Section 3 elaborates on the method and model for the analysis of road transport CO_2 emissions. Section 4 presents several policy scenarios and related parameters. Section 5 analyzes and discusses the research results of this study. Section 6 summarizes the main findings and makes policy suggestions.

2. Literature Review

2.1. Calculation of Road Transport CO_2 Emissions

Given the complexity and dispersion of transportation CO_2 emissions, there are still great challenges in the quantitative assessment of road transport CO_2 emissions. Referring to the national greenhouse gas inventory guidelines of the Intergovernmental Panel on Climate Change (IPCC), the current calculation methods of CO_2 emissions from mobile sources mainly include the "top-down" and "bottom-up" approaches [4]. The "top-down" method calculates the road transport CO_2 emissions by multiplying the total consumption of various fuels by their corresponding CO_2 emission factors [5]. For the "bottom-up" method, the vehicle mileages, fuel structures, and fuel economy of various types of vehicles should be collected to calculate the total CO_2 emissions of road transport [6]. For comparison, the "top-down" method has advantages in data collection and calculation, while the "bottom-up" method provides more details about the contribution of CO_2 emissions. In practice, researchers usually combine both approaches to calculate CO_2 emissions from road transport [7]. For a life cycle assessment of road transport CO_2 emissions, the "life cycle" method should be used to calculate the total CO_2 emissions of upstream, midstream, and downstream industries associated with road transport, including the production, transportation, use, disposal of vehicle, and fuels [8]. Due to the lack of data on private transport, most of the existing empirical studies of road transport CO_2 emissions focus on operational transportation, such as buses, taxis, intercity passenger transport, intercity freight transport, etc. To fill this gap, this study accounts for both operational and private transport and analyzes the CO_2 emission reduction potential of different road transport subsectors.

2.2. Influencing Factors of Road Transport CO_2 Emissions

Presently, many scholars are committed to studying the influencing factors of road transport CO_2 emissions and have produced more systematic research results in this field. For example, Zhang et al. [9] divided the main factors affecting transport carbon dioxide emissions into three categories: demand-side factors, supply-side factors, and environmental measurement factors. To further analyze the impact of different factors on road transport CO_2 emissions, the CO_2 emission contribution of each factor can be obtained by the Log-Mean Divisia Index (LMDI) method, and targeted policy suggestions can be presented. Based on this method, Liu et al. [10] investigated the main factors of CO_2 emissions in the field of transportation and formulated corresponding energy-saving policies in China. Quan et al. [11] analyzed the influences of the CO_2 emission factor, energy intensity, energy structure, economic level, and population on CO_2 emissions from China's logistics industry. However, most of the existing studies neglect the influence of power generation. With the substantial increase in the market share of electric vehicles in the future, the emission factor of power generation will also become a key factor affecting the CO_2 emissions of road transportation. Therefore, this study will explore the contribution of clean electricity to the CO_2 emission reduction of road transport.

2.3. Emission Reduction Potential of Low-Carbon Transport Policy

The formulation and implementation of low-carbon transport policies are the main ways to promote the low-carbon transition of transport. Many studies analyze the road transport emission reduction potential under different low-carbon transport scenarios. In these studies, the impacts of various policies and combinations, such as fuel tax, fuel labeling, new car purchase tax reduction, high emission vehicle purchase penalty tax, vehicle scrapping incentive, and vehicle transport restriction, on transport CO_2 emissions have been widely analyzed and discussed [12–21]. The bottom-up models, such as the Transport and Mobility Leuven (TREMOVE), Integrated Land Use and Transport Modeling System (TRANUS), Integrated MARKAL-EFOM System (TIMES), and Low Emissions Analysis Platform (LEAP), can be used to evaluate the emission reduction potential of low-carbon transport policies. Local parameters such as vehicle population, vehicle structure, fuel economy, and emission factors are usually inputted into the model [22–24]. Among them, the LEAP is one of the most popular analytical methods because it requires less data and can provide a comprehensive scenario analysis. For example, based on the LEAP, Feng et al. [25] analyzed the trend of energy demand, pollutants, and carbon emissions in China's transportation sector under three policy scenarios: pollution reduction (PR), low carbon (LC), and the deep-seated low carbon (DLC) scenarios. Pang et al. [26] analyzed the impact of three scenarios with different policies and measures on greenhouse gas emissions from road transport in Lanzhou, China, from 2015 to 2040 based on the LEAP. However, most scenarios in the existing studies did not include certain future policy scenarios, such as electric vehicles with cleaner power, shared autonomous vehicles, CO_2 emission trading markets, etc. Therefore, the emission reduction potential of road transport may be underestimated [27]. To help reach the goal of carbon neutrality in China, this study will propose more emerging and comprehensive policy scenarios and assess the CO_2 emission reduction potential of road transport under these scenarios.

3. Methods

3.1. Research Framework

To evaluate the CO_2 emission reduction potential of China's road transport sector and the contribution of various influencing factors, this study integrates the LEAP with the LMDI method. The research framework is shown in Figure 1. First, the future population of China is predicted based on the death rate and birth rate of the historical population. Meanwhile, the future vehicle population is predicted based on the Gompertz model. Second, we set eight policy scenarios with different parameters, such as vehicle structure, vehicle mileage, fuel economy, and emission factors. Third, the LEAP is used to predict CO_2 emissions of road transport in China from 2020 to 2060 and analyze the CO_2 emission reduction potential of different scenarios. Last, the LMDI method is applied to calculate the contribution of various factors to CO_2 emission reduction in each scenario.

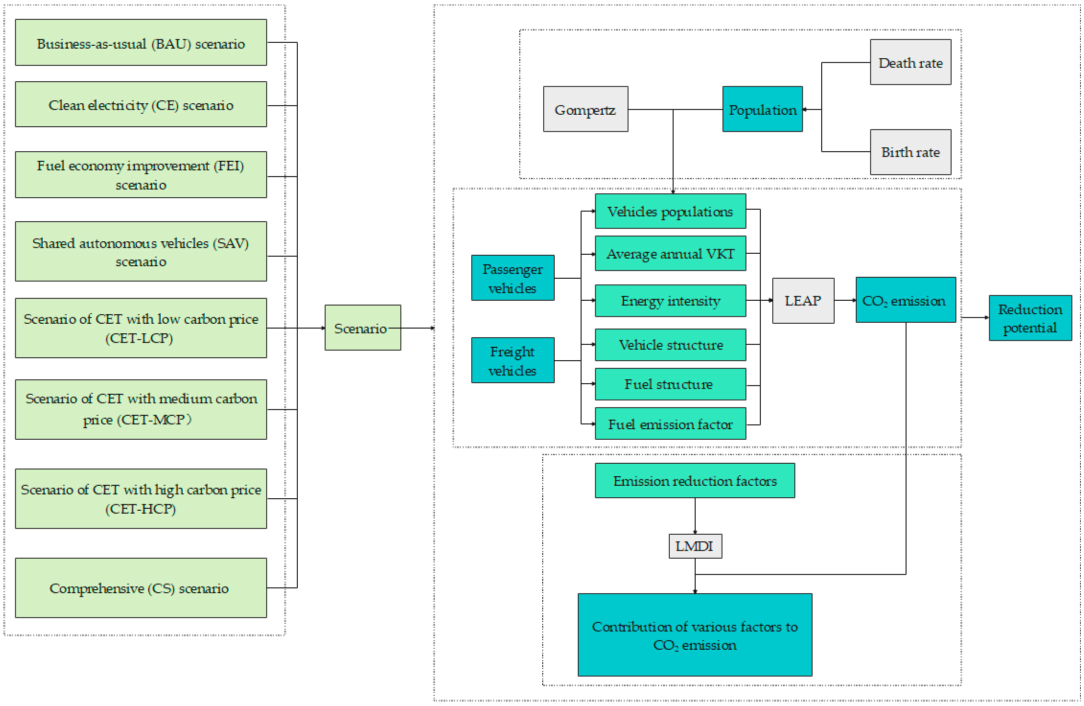


Figure 1. Research framework.

3.2. Road Transport CO_2 Emission Calculation

In this study, the CO_2 emissions of road transport are calculated by the bottom-up method. This method calculates road transport CO_2 emissions based on the "ASIF" framework [2,3], where "A" represents the travel activity, "S" represents the mode share, "I" represents the energy intensity of each mode, fuel, and vehicle type, and "F" represents and carbon content of each fuel to total emissions. The road transport CO_2 emissions include emissions from both passenger and freight transport. The equation is expressed as follows:

$$C = C_P + C_T \tag{1}$$

where C_P is the total CO_2 emissions of road passenger transport, C_T is the total CO_2 emissions of road freight transport, and C is the total CO_2 emissions of road transport.

The equation for calculating passenger transport CO_2 emissions is expressed as follows:

$$C_P = \sum_{i=1}^{4} \sum_{k=1}^{3} N_P S_i M_i Q_{ik} F_{ik} E_k \tag{2}$$

where i = 1, 2, 3, and 4 refer to mini, small, medium, and large passenger vehicle types, respectively; k = 1, 2, and 3 represent fuel types of gasoline, electricity, and hybrid, respectively; N_P is the total number of passenger vehicles; S_i is the proportion of passenger vehicles of type i; M_i is the average annual VKT of vehicles of type i; Q_{ik} is the proportion of fuels of type k for vehicles of type i; F_{ik} is the fuel economy of vehicles of type i with fuel type k; and E_k is the CO_2 emission factor of fuel of type k.

The equation for calculating freight transport CO_2 emissions is expressed as follows:

$$C_T = \sum_{j=1}^{4} \sum_{h=1}^{4} N_T S_j M_j Q_{jh} F_{jh} E_h \tag{3}$$

where j = 1, 2, 3, 4 refer to mini, light, medium, and heavy freight vehicles, respectively; h = 1, 2, 3, 4 represent fuel types of diesel, gasoline, electricity, and natural gas, respectively; N_t is the total number of freight vehicles; S_j is the proportion of freight vehicles of type j; M_j is the average annual VKT of trucks of type j; Q_{jh} is the proportion of fuels of type h for freight vehicles of type j; F_{jh} is the fuel economy of trucks of type j with fuel type h; and E_h is the CO_2 emission factor of fuel of type h.

According to Equations (1)–(3), Equation (1) is equivalent to:

$$C = \sum_{i=1}^{4}\sum_{k=1}^{3} N_P S_i M_i Q_{ik} F_{ik} E_k + \sum_{j=1}^{4}\sum_{h=1}^{4} N_T S_j M_j Q_{jh} F_{jh} E_h \quad (4)$$

3.3. Low Emissions Analysis Platform (LEAP)

The LEAP is mainly aimed at the whole process of terminal energy consumption and comprehensively evaluates the influence of various technologies and policies on energy conservation and emission reduction from the aspects of energy supply structure, energy technology level, energy demand, etc., which is more consistent with the content and goal of research on the low-carbon development path of urban transportation [28–30]. The LEAP is a medium- and long-term modeling tool. In most studies that use the LEAP, the prediction period is generally 20 to 50 years. Therefore, this study uses the LEAP to predict the changing trend of China's road transport CO_2 emissions in different policy scenarios from 2020 to 2060. The parameter settings of these scenarios are introduced in the next section. We can then analyze the characteristics of peak CO_2 emissions years and trends in different scenarios and compares them with the BAU scenario.

3.4. Log-Mean Divisia Index (LMDI) Method

The LMDI method is a complete decomposition analysis method without residual error [31]. Using the LMDI method, the contribution of different factors to CO_2 emission reduction can be examined. This study decomposes the change in road transport CO_2 emissions from base year 0 to target year t into six parts of contribution, as shown in Equation (5).

$$\Delta C = C^t - C^0 = \Delta C_N + \Delta C_S + \Delta C_Q + \Delta C_F + \Delta C_M + \Delta C_E \quad (5)$$

where C^t represents the CO_2 emissions of road transport in the target year; C^0 represents the CO_2 emissions of road transport in the base year; ΔC_N represents the contribution of vehicle population; ΔC_S represents the contribution of vehicle structure; ΔC_Q represents the contribution of fuel structure; ΔC_F represents the contribution of fuel economy; ΔC_M represents the contribution of average annual VKT, and ΔC_E represents the contribution of fuel CO_2 emission factors.

According to the decomposition method of the LMDI, each item on the right side of Equation (5) can be expressed as follows:

$$\Delta C_N = \sum_{i=1}^{4} \frac{C_i^t - C_i^0}{\ln C_i^t - \ln C_i^0} \ln\left(\frac{N_P^t}{N_P^0}\right) + \sum_{j=1}^{4} \frac{C_j^t - C_j^0}{\ln C_j^t - \ln C_j^0} \ln\left(\frac{N_T^t}{N_T^0}\right) \quad (6)$$

$$\Delta C_S = \sum_{i=1}^{4} \frac{C_i^t - C_i^0}{\ln C_i^t - \ln C_i^0} \ln\left(\frac{S_i^t}{S_i^0}\right) + \sum_{j=1}^{4} \frac{C_j^t - C_j^0}{\ln C_j^t - \ln C_j^0} \ln\left(\frac{S_j^t}{S_j^0}\right) \quad (7)$$

$$\Delta C_Q = \sum_{i=1}^{4}\sum_{k=1}^{3} \frac{C_{ik}^t - C_{ik}^0}{\ln C_{ik}^t - \ln C_{ik}^0} \ln\left(\frac{Q_{ik}^t}{Q_{ik}^0}\right) + \sum_{j=1}^{4}\sum_{h=1}^{4} \frac{C_{jh}^t - C_{jh}^0}{\ln C_{jh}^t - \ln C_{jh}^0} \ln\left(\frac{Q_{jh}^t}{Q_{jh}^0}\right) \quad (8)$$

$$\Delta C_F = \sum_{i=1}^{4}\sum_{k=1}^{3} \frac{C_{ik}^t - C_{ik}^0}{\ln C_{ik}^t - \ln C_{ik}^0} \ln\left(\frac{F_{ik}^t}{F_{ik}^0}\right) + \sum_{j=1}^{4}\sum_{h=1}^{4} \frac{C_{jh}^t - C_{jh}^0}{\ln C_{jh}^t - \ln C_{jh}^0} \ln\left(\frac{F_{jh}^t}{F_{jh}^0}\right) \quad (9)$$

$$\Delta C_M = \sum_{i=1}^{4} \frac{C_i^t - C_i^0}{\ln C_i^t - \ln C_i^0} \ln\left(\frac{M_i^t}{M_i^0}\right) + \sum_{j=1}^{4} \frac{C_j^t - C_j^0}{\ln C_j^t - \ln C_j^0} \ln\left(\frac{M_j^t}{M_j^0}\right) \quad (10)$$

$$\Delta C_E = \sum_{i=1}^{4}\sum_{k=1}^{3}\frac{C_{ik}^t - C_{ik}^0}{\ln C_{ik}^t - \ln C_{ik}^0}\ln(\frac{E_k^t}{E_k^0}) + \sum_{j=1}^{4}\sum_{h=1}^{4}\frac{C_{jh}^t - C_{jh}^0}{\ln C_{jh}^t - \ln C_{jh}^0}\ln(\frac{E_h^t}{E_h^0}) \quad (11)$$

To further compare and analyze the relative contribution of each factor under different scenarios, this study calculates the contribution rate of each factor to the total change of road transport CO_2 emissions as follows:

$$CR(l) = \Delta C_l / \Delta C \quad (12)$$

where $CR(l)$ is the contribution rate of factor l and ΔC_l is the CO_2 emission contribution of factor l. If $CR(l) > 0$, then the influencing factor l may drive the increase of CO_2 emissions. Otherwise, it contributes to reducing CO_2 emissions.

4. Scenario Setting

This study aims to analyze the CO_2 emission reduction potential of road transport from 2020 to 2060. Therefore, different policy scenarios, such as the business-as-usual (BAU), clean electricity (CE), fuel economy improvement (FEI), shared autonomous vehicles (SAV), CO_2 emission trading (CET) (with low, medium and high carbon prices), and comprehensive (CS) scenarios, are established based on the key influencing factors of road transport CO_2 emissions. The relevant scenario parameters involved in this study include vehicle structure, fuel economy, fuel emission factor, average annual VKT of vehicles, etc. The settings of these scenarios are presented hereafter.

4.1. Business-as-Usual (BAU) Scenario

In the BAU scenario, we assume the CO_2 emission of road transport keeps the historical development trend with no additional policy implemented in the future. According to the national statistical yearbook, China's total population in 2020 was 1.412 billion, with a birth rate of 8.52‰ and a death rate of 7.09‰. This research set the initial death rate DR(0) = 7.09‰, which increases yearly with an increment of 0.17‰, and the initial birth rate BR(0) = 8.52‰, which decreases yearly with a decrease of 0.2‰. The projection of the population of China from 2010 to 2060 in the BAU scenario is shown in Figure 2.

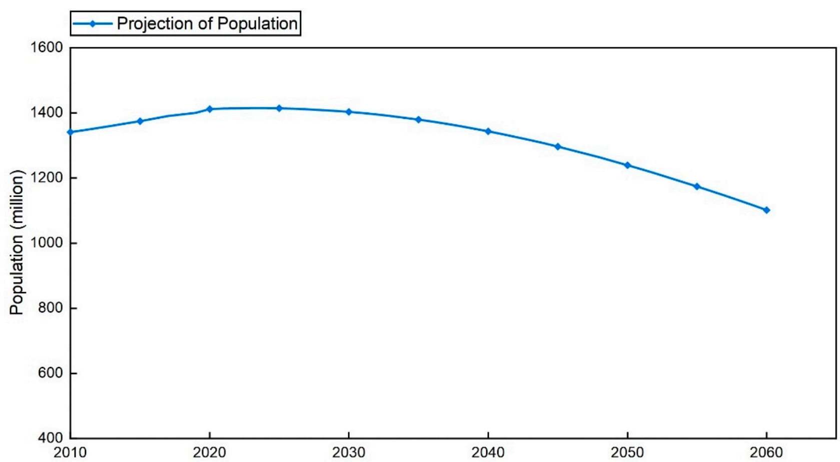


Figure 2. Projection of the population of China from 2010 to 2060 in the BAU scenario.

Based on the forecast of the International Energy Agency [32–34], the car ownership per thousand people in China will be approximately 494 in 2050. Therefore, the number of vehicles per 1000 people in China in the BAU scenario in 2060 is set to 500 cars per 1000 people [35]. The Gompertz model is then used to predict the vehicle population in China from 2020 to 2060 based on historical data from 2010 to 2020. The projections of China's vehicle population, freight vehicle population, and passenger vehicle population from 2010 to 2060 are shown in Figure 3.

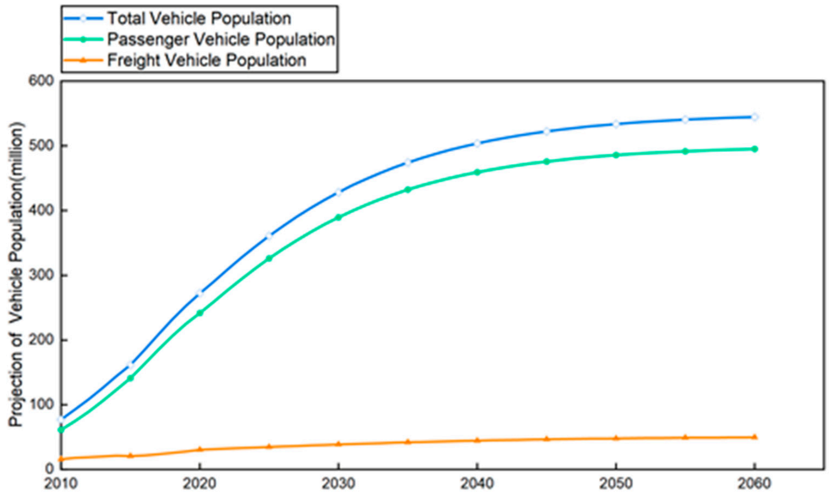


Figure 3. Projection of the vehicle population of China from 2010 to 2060 in the BAU scenario.

Based on historical data and related studies [36–41], we predict the vehicle structure, fuel structure, fuel economy, average annual VKT, and fuel emission factor of road transport from 2020 to 2060 for the BAU scenario. The settings of these parameters for passenger transport and freight transport in BAU scenarios are summarized in Tables 1 and 2.

Table 1. Parameter settings of passenger transport in the BAU scenario.

Influencing Factors of CO_2 Emission	Vehicle Type	Fuel Type	2020	2030	2040	2050	2060
Passenger vehicle population (ten thousand)	All	All	24,166.2	38,958.2	45,927.4	48,574.4	49,510.0
Vehicle structure	Mini	All	0.65%	0.10%	0.01%	0.00%	0.00%
	Small	All	98.40%	99.13%	99.22%	99.23%	99.23%
	Medium	All	0.28%	0.38%	0.52%	0.58%	0.62%
	Large	All	0.67%	0.38%	0.26%	0.19%	0.15%
Fuel structure	Mini, Small, and Medium	Gasoline	98%	86%	68%	34%	0%
		Electricity	1.6%	12.74%	30.72%	64.68%	100%
		Hybrid	0.4%	1.26%	1.28%	1.32%	0%
	Large	Gasoline	75%	50%	30%	15%	0%
		Electricity	25%	50%	70%	85%	100%
Fuel economy	Mini	Gasoline (L/100 km)	5.20	4.03	3.84	3.67	3.53
		Electricity (kWh/100 km)	8.70	6.73	6.40	6.13	5.89
		Hybrid (L/100 km)	2.81	2.18	2.07	1.98	1.91
	Small	Gasoline (L/100 km)	8.30	7.50	6.75	6.00	5.30
		Electricity (kWh/100 km)	13.00	10.00	9.00	8.20	7.80
		Hybrid (L/100 km)	4.22	3.36	2.75	2.60	2.58
	Medium	Gasoline (L/100 km)	17.10	15.20	14.80	14.60	14.50
		Electricity (kWh/100 km)	120.00	116.00	110.00	100.00	92.00
		Hybrid (L/100 km)	9.24	8.22	8.00	7.89	7.84
	Large	Gasoline (L/100 km)	21.80	19.40	18.90	18.50	18.20
		Electricity (kWh/100 km)	144.00	140.00	135.0	128.00	122.00
Average annual VKT (km)	Mini	All	10,000	8917	7917	7000	6500
	Small	All	12,000	10,700	9500	8500	7500
	Medium	All	35,000	35,750	36,500	37,250	38,000
	Large	All	48,300	48,900	49,200	49,500	49,800
Fuel emission factor	All	Gasoline (kg/L)	2.42	2.42	2.42	2.42	2.42
	All	Electricity (kg/kWh)	0.71	0.52	0.46	0.4	0.38

Table 2. Parameter settings of freight transport in the BAU scenario.

Influencing Factors of CO_2 Emission	Vehicle Type	Fuel Type	2020	2030	2040	2050	2060
Freight vehicle population (ten thousand)	All	All	3042.6	3861.4	4453.5	4793.3	4957.4
Vehicle structure	Mini	Diesel	0.00%	0.00%	0.00%	0.00%	0.00%
	Light		40.20%	37.68%	33.75%	29.12%	24.50%
	Medium		5.70%	3.00%	1.70%	1.10%	0.50%
	Heavy		54.10%	59.33%	64.55%	69.78%	75.00%
	Mini	Gasoline	5.50%	4.00%	2.40%	1.20%	0.00%
	Light		94.50%	96.00%	97.60%	98.80%	100.00%
	Medium		0.00%	0.00%	0.00%	0.00%	0.00%
	Heavy		0.00%	0.00%	0.00%	0.00%	0.00%
	Mini	Electricity	99.50%	83.38%	67.25%	51.13%	35.00%
	Light		0.50%	7.88%	15.25%	22.63%	30.00%
	Medium		0.00%	5.00%	10.00%	15.00%	20.00%
	Heavy		0.00%	3.75%	7.50%	11.25%	15.00%
	Mini	Natural gas	0.10%	0.00%	0.00%	0.00%	0.00%
	Light		3.80%	2.50%	1.60%	0.80%	0.00%
	Medium		0.40%	0.00%	0.00%	0.00%	0.00%
	Heavy		95.70%	97.50%	98.40%	99.20%	100.00%
Fuel structure	All	Diesel	69.60%	60.00%	51.00%	43.00%	35.00%
	All	Gasoline	27.90%	19.30%	14.00%	12.00%	10.00%
	All	Electricity	0.70%	13.40%	24.00%	32.00%	40.00%
	All	Natural gas	1.80%	7.30%	11.00%	13.00%	15.00%
Fuel Economy	Mini	Diesel (L/100 km)	6.80	6.10	5.80	5.60	5.50
		Gasoline (L/100 km)	9.60	8.60	7.40	5.80	4.20
		Natural gas (m^3/100 km)	8.40	7.50	6.68	5.84	5.00
		Electricity (kWh/100 km)	18.50	17.40	16.00	15.20	14.70
	Light	Diesel (L/100 km)	8.70	7.80	7.40	7.10	7.00
		Gasoline (L/100 km)	11.00	9.90	8.84	7.72	6.60
		Natural gas (m^3/100 km)	11.20	10.10	8.94	7.82	6.70
		Electricity (kWh/100 km)	125.00	119.00	113.00	106.00	102.00
	Medium	Diesel (L/100 km)	15.50	14.70	14.00	13.40	12.90
		Natural gas (m^3/100 km)	17.50	15.70	14.02	12.26	10.50
		Electricity (kWh/100 km)	132.00	128.00	123.00	114.00	111.00
	Heavy	Diesel (L/100 km)	32.60	30.80	29.30	28.00	27.00
		Natural gas (m^3/100 km)	30.80	27.80	24.66	21.58	18.50
		Electricity (kWh/100 km)	150.00	146.00	140.00	132.00	129.00
Average annual VMT (km)	Mini	All	20,000	20,000	20,000	20,000	20,000
	Light	All	20,000	20,000	20,000	20,000	20,000
	Medium	All	24,000	25,627	27,498	29,288	31,000
	Heavy	All	40,000	40,500	41,143	41,786	42,500
Fuel emission factor	All	Diesel (kg/L)	2.8	2.8	2.8	2.8	2.8
	All	Gasoline (kg/L)	2.42	2.42	2.42	2.42	2.42
	All	Natural gas (kg/m^3)	2.62	2.62	2.62	2.62	2.62
	All	Electricity (kg/kWh)	0.71	0.52	0.46	0.4	0.38

4.2. Clean Electricity (CE) Scenario

Presently, China's energy structure is dominated by coal. Therefore, the power of electric vehicles is mainly generated from coal and the CO_2 emissions of electric vehicles are still high. With the diffusion of electric vehicles in the future, the influence of the electricity emission factor on road transport CO_2 emission will be more significant. This study proposes a clean electricity (CE) scenario where renewable energy power generation will be rapidly developed and widely adopted. In the CE scenario, the CO_2 emission factor of electricity should decrease more significantly compared with the BAU scenario. Therefore, the CO_2 emission factors of electricity in the CE scenario are assumed to decrease 2.5 times as fast as in the BAU scenario. The specific parameters that change compared with the BAU scenario are shown in Table 3. The rest of the parameters are the same as those in the BAU scenario.

Table 3. Parameter settings in the CE scenario.

Influencing Factors of CO_2 Emission	2020	2030	2040	2050	2060
CO_2 emission factor of electricity (kg/kWh)	0.71	0.32	0.23	0.16	0.14

4.3. Fuel Economy Improvement (FEI) Scenario

As one of the main factors affecting CO_2 emissions, the fuel economy of vehicles will be improved with the continuous development of automobile conservation technology in the future. Therefore, we propose a fuel economy improvement (FEI) scenario where the fuel economy of vehicles will be increased faster than that in the BAU scenario. Due to the differences in power types and engine technology, the room for improvement in the fuel economy of freight vehicles is usually larger than that of passenger vehicles. Thus, we set the improvement rate of fuel economy for passenger vehicles and freight vehicles in the FEI scenario to be 1.5 times and 1.8 times, respectively, that in the BAU scenario. The specific parameters that change compared with the BAU scenario are shown in Table 4. The rest of the parameters are the same as those in the BAU scenario.

Table 4. Parameter settings in the FEI scenario.

Influencing Factors of CO_2 Emission	Vehicle Type		Fuel Type	2020	2030	2040	2050	2060
Fuel economy	Passenger vehicle	Mini	Gasoline (L/100 km)	5.20	3.54	3.28	3.07	2.89
			Electricity (kWh/100 km)	8.70	5.90	5.48	5.12	4.83
			Hybrid (L/100 km)	2.81	1.91	1.77	1.66	1.56
		Small	Gasoline (L/100 km)	8.30	7.13	6.08	5.09	4.23
			Electricity (kWh/100 km)	13.00	8.75	7.47	6.49	6.02
			Hybrid (L/100 km)	4.22	2.99	2.21	2.03	2.01
		Medium	Gasoline (L/100 km)	17.10	14.32	13.76	13.48	13.34
			Electricity (kWh/100 km)	120.00	114.05	105.30	91.24	80.49
			Hybrid (L/100 km)	9.24	7.74	7.44	7.29	7.21
		Large	Gasoline (L/100 km)	21.80	18.29	17.59	17.03	16.62
			Electricity (kWh/100 km)	144.00	138.04	130.70	120.66	112.26
	Freight vehicle	Mini	Diesel (L/100 km)	6.80	5.59	5.10	4.79	4.64
			Gasoline (L/100 km)	9.60	7.87	6.00	3.85	2.14
			Natural gas (m^3/ 100 km)	8.40	6.84	5.55	4.35	3.28
			Electricity (kWh/100 km)	18.50	16.56	14.23	12.98	12.22
		Light	Diesel (L/100 km)	8.70	7.14	6.49	6.03	5.88
			Gasoline (L/100 km)	11.00	9.09	7.41	5.80	4.36
			Natural gas (m^3/100 km)	11.20	9.29	7.45	5.84	4.42
			Electricity (kWh/100 km)	125.00	114.39	104.20	92.84	86.62
		Medium	Diesel (L/100 km)	15.50	14.09	12.90	11.92	11.13
			Natural gas (m^3/100 km)	17.50	14.37	11.72	9.19	6.94
			Electricity (kWh/100 km)	132.00	124.88	116.33	101.32	96.57
		Heavy	Diesel (L/100 km)	32.60	29.43	26.89	24.78	23.21
			Natural gas (m^3/100 km)	30.80	25.59	20.60	16.18	12.24
			Electricity (kWh/100 km)	150.00	142.87	132.46	119.12	114.28

4.4. Shared Autonomous Vehicles (SAV) Scenario

The emerging shared autonomous vehicles may steer a revolution in passenger transport. According to related research [42–45], people will give up purchasing private cars if shared mobility and autonomous driving services are widely adopted in the future. Since the shared autonomous vehicles usually have high capacity, the vehicle structure of passenger transport will change significantly. In addition, shared autonomous vehicles can also improve transport efficiency and fuel economy of vehicles. Therefore, we propose a shared autonomous vehicles (SAV) scenario where the vehicle population of passenger vehicles, the structure of passenger vehicles, and the average annual VKT of passenger vehicles are changed compared with the BAU scenario. In the SAV scenario, the car ownership per 1000 people is set to 400 in 2060, which is less than that in the BAU scenario. Besides, the proportion of medium and large passenger vehicles is greater than that in the BAU scenario. The proportion of medium and large passenger vehicles is assumed to increase 0.5 times as fast as in the BAU scenario. In addition, we set the improvement rate of fuel economy for vehicles in the SAV scenario to be 1.25 times that in the BAU scenario. The specific parameters that change compared with the BAU scenario are shown in Table 5. The rest of the parameters are the same as those in the BAU scenario.

Table 5. Parameter settings in the SAV scenario.

Influencing Factors of CO_2 Emission	Vehicle Type		Fuel Type	2020	2030	2040	2050	2060
Passenger vehicle population (ten thousand)	All		All	24,166.2	34,848.5	38,095.5	38,854.2	38,991.0
Vehicle structure	Passenger vehicle	Mini	All	0.65%	0.30%	0.13%	0.06%	0.03%
		Small	All	98.40%	98.77%	98.81%	98.81%	98.81%
		Medium	All	0.28%	0.43%	0.65%	0.77%	0.84%
		Large	All	0.67%	0.50%	0.41%	0.35%	0.32%
Fuel economy	Passenger vehicle	Mini	Gasoline (L/100 km)	5.20	3.78	3.55	3.36	3.19
			Electricity (kWh/100 km)	8.70	6.31	5.92	5.61	5.33
			Hybrid (L/100 km)	2.81	2.04	1.92	1.82	1.73
		Small	Gasoline (L/100 km)	8.30	7.31	6.41	5.53	4.73
			Electricity (kWh/100 km)	13.00	9.35	8.20	7.30	6.85
			Hybrid (L/100 km)	4.22	3.17	2.46	2.30	2.28
		Medium	Gasoline (L/100 km)	17.10	14.76	14.27	14.03	13.91
			Electricity (kWh/100 km)	120.00	115.02	107.63	95.52	86.06
			Hybrid (L/100 km)	9.24	7.98	7.71	7.58	7.52
		Large	Gasoline (L/100 km)	21.80	18.84	18.23	17.75	17.39
			Electricity (kWh/100 km)	144.00	139.02	132.83	124.28	117.03
	Freight vehicle	Mini	Diesel (L/100 km)	6.80	5.94	5.57	5.33	5.21
			Gasoline (L/100 km)	9.60	8.36	6.93	5.11	3.40
			Natural gas (m^3/100 km)	8.40	7.29	6.31	5.33	4.39
			Electricity (kWh/100 km)	18.50	17.13	15.43	14.47	13.88
		Light	Diesel (L/100 km)	8.70	7.59	7.10	6.75	6.63
			Gasoline (L/100 km)	11.00	9.64	8.37	7.06	5.80
			Natural gas (m^3/100 km)	11.20	9.84	8.45	7.14	5.89
			Electricity (kWh/100 km)	125.00	117.54	110.18	101.70	96.93
		Medium	Diesel (L/100 km)	15.50	14.51	13.65	12.92	12.32
			Natural gas (m^3/100 km)	17.50	15.28	13.26	11.21	9.23
			Electricity (kWh/100 km)	132.00	127.02	120.84	109.88	106.28
		Heavy	Diesel (L/100 km)	32.6	30.36	28.53	26.95	25.75
			Natural gas (m^3/100 km)	30.8	27.09	23.32	19.73	16.27
			Electricity (kWh/100 km)	150.00	145.02	137.60	127.84	124.21

4.5. CO_2 Emission Trading (CET) Scenario

CO_2 emissions trading is a cost-effective climate policy to reduce greenhouse gas emissions. Although the road transport sector has not currently been incorporated into the emission trading system, it is very likely to be implemented in the future with the development of emerging technology and the maturity of the carbon market. This study proposes a CO_2 emissions trading (CET) scenario with different levels of carbon prices (low, mid, and high). According to related research [46], the carbon price of CET can stimulate vehicle users to choose low-carbon vehicles (with fuel types of electricity, hybrid, and natural gas) and decrease the average annual VKT. The higher the price, the greater the influences.

In the CO_2 emissions trading scenario with low carbon prices (CET-LCP), the proportion of low-carbon vehicles is set to 1.25 times that in the BAU scenario and the average annual VKT of vehicles in 2060 is set to 0.85 times that in the BAU scenario. In the CO_2 emissions trading scenario with mid carbon prices (CET-MCP), the proportion of low-carbon vehicles is set to 1.5 times that in the BAU scenario and the average annual VKT of vehicles in 2060 is set to 0.75 times that in the BAU scenario. In the CO_2 emissions trading scenario with high carbon prices (CET-HCP), the proportion of low-carbon vehicles is set to 2 times that in the BAU scenario and the average annual VKT of vehicles in 2060 is set to 0.65 times that in the BAU scenario. The specific parameters of the three scenarios that change compared with the BAU scenario are shown in Tables 6–8. The rest of the parameters are the same as those in the BAU scenario.

Table 6. Parameter settings in the CET-LCP scenario.

Influencing Factors of CO_2 Emission	Vehicle Type		Fuel Type	2020	2030	2040	2050	2060
Fuel structure	Passenger vehicle	Small and medium	Gasoline	98.00%	82.50%	60.00%	17.50%	0.00%
			Electricity	1.60%	15.93%	38.40%	80.85%	100.00%
			Hybrid	0.40%	1.58%	1.60%	1.65%	0.00%
		Large	Gasoline	75.00%	37.50%	12.50%	0.00%	0.00%
			Electricity	25.00%	62.50%	87.50%	100.00%	100.00%
	Freight vehicle		Diesel	69.60%	56.08%	44.13%	34.20%	24.31%
			Gasoline	27.90%	18.04%	12.12%	9.55%	6.94%
			Natural gas	1.80%	9.13%	13.75%	16.25%	18.75%
			Electricity	0.70%	16.75%	30.00%	40.00%	50.00%
Average annual VMT (km)	Passenger vehicle	Mini	All	10,000	8881	7763	6644	5525
		Small	All	12,000	10,594	9188	7781	6375
		Medium	All	35,000	34,325	33,650	32,975	32,300
		Large	All	48,300	46,808	45,315	43,823	42,330
	Freight vehicle	Mini	All	20,000	19,250	18,500	17,750	17,000
		Light	All	20,000	19,250	18,500	17,750	17,000
		Medium	All	24,000	24,588	25,175	25,763	26,350
		Heavy	All	40,000	39,031	38,063	37,094	36,125

Table 7. Parameter settings in the CET-MCP scenario.

Influencing Factors of CO_2 Emission	Vehicle Type		Fuel Type	2020	2030	2040	2050	2060
Fuel structure	Passenger vehicle	Small and medium	Gasoline	98.00%	79.00%	52.00%	1.00%	0.00%
			Electricity	1.60%	19.11%	46.08%	97.02%	100.00%
			Hybrid	0.40%	1.89%	1.92%	1.98%	0.00%
		Large	Gasoline	75.00%	25.00%	0.00%	0.00%	0.00%
			Electricity	25.00%	75.00%	100.00%	100.00%	100.00%
	Freight vehicle		Diesel	69.60%	52.17%	37.27%	25.41%	13.61%
			Gasoline	27.90%	16.78%	10.23%	7.09%	3.89%
			Natural gas	1.80%	10.95%	16.50%	19.50%	22.50%
			Electricity	0.70%	20.10%	36.00%	48.00%	60.00%

Table 7. *Cont.*

Influencing Factors of CO_2 Emission	Vehicle Type		Fuel Type	2020	2030	2040	2050	2060
Average annual VMT (km)	Passenger vehicle	Mini	All	10,000	8719	7438	6156	4875
		Small	All	12,000	10,406	8813	7219	5625
		Medium	All	35,000	33,375	31,750	30,125	28,500
		Large	All	48,300	45,563	42,825	40,088	37,350
	Freight vehicle	Mini	All	20,000	18,750	17,500	16,250	15,000
		Light	All	20,000	18,750	17,500	16,250	15,000
		Medium	All	24,000	23,813	23,625	23,438	23,250
		Heavy	All	40,000	37,969	35,938	33,906	31,875

Table 8. Parameter settings in the CET-HCP scenario.

Influencing Factors of CO_2 Emission	Vehicle Type		Fuel Type	2020	2030	2040	2050	2060
Fuel structure	Passenger vehicle	Small and medium	Gasoline	98.00%	74.80%	42.40%	0.00%	0.00%
			Electricity	1.60%	22.93%	55.30%	98%	100.00%
			Hybrid	0.40%	2.27%	2.30%	2.00%	0.00%
		Large	Gasoline	75.00%	0.00%	0.00%	0.00%	0.00%
			Electricity	25.00%	100.00%	100.00%	100.00%	100.00%
	Freight vehicle		Diesel	69.60%	44.34%	23.54%	7.82%	0.00%
			Gasoline	27.90%	14.26%	6.46%	2.18%	0.00%
			Natural gas	1.80%	14.60%	22.00%	26.00%	27.27%
			Electricity	0.70%	26.80%	48.00%	64.00%	72.73%
Average annual VMT (km)	Passenger vehicle	Mini	All	10,000	8556	7113	5669	4225
		Small	All	12,000	10,219	8438	6656	4875
		Medium	All	35,000	32,425	29,850	27,275	24,700
		Large	All	48,300	44,318	40,335	36,353	32,370
	Freight vehicle	Mini	All	20,000	18,250	16,500	14,750	13,000
		Light	All	20,000	18,250	16,500	14,750	13,000
		Medium	All	24,000	23,038	22,075	21,113	20,150
		Heavy	All	40,000	36,906	33,813	30,719	27,625

4.6. Comprehensive (CS) Scenario

The CS scenario is the combination of the CE, SAV, FEI, and CET-HCP scenarios proposed above. In this scenario, all six factors affecting road transport CO_2 emissions—vehicle population, vehicle structure, fuel structure, fuel emission factor, fuel economy, and average annual VKT—are changed. The specific parameters that change compared with the BAU scenario are shown in Tables 9 and 10. The rest of the parameters are the same as those in the BAU scenario.

Table 9. Parameter settings of passenger vehicles in the CS scenario.

Influencing Factors of CO_2 Emission	Vehicle Type	Fuel Type	2020	2030	2040	2050	2060
Passenger vehicles population (ten thousand)	All	All	3042.6	5133.3	6864.9	7512.9	8731.0
Vehicle structure	Passenger vehicle	Mini	All	0.65%	0.30%	0.13%	0.06%
		Small	All	98.40%	98.77%	98.81%	98.81%
		Medium	All	0.28%	0.43%	0.65%	0.77%
		Large	All	0.67%	0.50%	0.41%	0.35%
Vehicle structure	Mini	All	0.65%	0.04%	0.00%	0.00%	0.00%
	Small	All	98.40%	98.77%	98.81%	98.81%	98.81%
	Medium	All	0.28%	0.69%	0.78%	0.83%	0.86%
	Large	All	0.67%	0.51%	0.42%	0.36%	0.32%

Table 9. *Cont.*

Influencing Factors of CO_2 Emission	Vehicle Type	Fuel Type	2020	2030	2040	2050	2060
Fuel structure	Small and medium	Gasoline	98.00%	74.80%	42.40%	0.00%	0.00%
		Electricity	1.60%	22.93%	55.30%	98%	100.00%
		Hybrid	0.40%	2.27%	2.30%	2.00%	0.00%
	Large	Gasoline	75.00%	0.00%	0.00%	0.00%	0.00%
		Electricity	25.00%	100.00%	100.00%	100.00%	100.00%
Fuel economy	Mini	Gasoline (L/100 km)	5.20	3.54	3.28	3.07	2.89
		Electricity (kWh/100 km)	8.70	5.90	5.48	5.12	4.83
		Hybrid (L/100 km)	2.81	1.91	1.77	1.66	1.56
	Small	Gasoline (L/100 km)	8.30	7.13	6.08	5.09	4.23
		Electricity (kWh/100 km)	13.00	8.75	7.47	6.49	6.02
		Hybrid (L/100 km)	4.22	2.99	2.21	2.03	2.01
	Medium	Gasoline (L/100 km)	17.10	14.32	13.76	13.48	13.34
		Electricity (kWh/100 km)	120.00	114.05	105.30	91.24	80.49
		Hybrid (L/100 km)	9.24	7.74	7.44	7.29	7.21
	Large	Gasoline (L/100 km)	21.80	18.29	17.59	17.03	16.62
		Electricity (kWh/100 km)	144.00	138.04	130.70	120.66	112.26
Average annual VMT (km)	Mini	All	10,000	8556	7113	5669	4225
	Small	All	12,000	10,219	8438	6656	4875
	Medium	All	35,000	32,425	29,850	27,275	24,700
	Large	All	48,300	44,318	40,335	36,353	32,370
Fuel emission factor	All	Electricity (kg/kWh)	0.71	0.32	0.23	0.16	0.14

Table 10. Parameter settings of freight vehicles in the CS scenario.

Influencing Factors of CO_2 Emission	Vehicle Type	Fuel Type	2020	2030	2040	2050	2060
Fuel structure	All	Diesel	69.60%	44.34%	23.54%	7.82%	0.00%
	All	Gasoline	27.90%	14.26%	6.46%	2.18%	0.00%
	All	Natural gas	1.80%	14.60%	22.00%	26.00%	27.27%
	All	Electricity	0.70%	26.80%	48.00%	64.00%	72.73%
Fuel economy	Mini	Diesel (L/100 km)	6.80	5.59	5.10	4.79	4.64
		Gasoline (L/100 km)	9.60	7.87	6.00	3.85	2.14
		Natural gas (m^3/100 km)	8.40	6.84	5.55	4.35	3.28
		Electricity (kWh/100 km)	18.50	16.56	14.23	12.98	12.22
	Light	Diesel (L/100 km)	8.70	7.14	6.49	6.03	5.88
		Gasoline (L/100 km)	11.00	9.09	7.41	5.80	4.36
		Natural gas (m^3/100 km)	11.20	9.29	7.45	5.84	4.42
		Electricity (kWh/100 km)	125.00	114.39	104.20	92.84	86.62
	Medium	Diesel (L/100 km)	15.50	14.09	12.90	11.92	11.13
		Natural gas (m^3/100 km)	17.50	14.37	11.72	9.19	6.94
		Electricity (kWh/100 km)	132.00	124.88	116.33	101.32	96.57
	Heavy	Diesel (L/100 km)	32.60	29.43	26.89	24.78	23.21
		Natural gas (m^3/100 km)	30.80	25.59	20.60	16.18	12.24
		Electricity (kWh/100 km)	150.00	142.87	132.46	119.12	114.28
Average annual VMT (km)	Mini	All	20,000	18,250	16,500	14,750	13,000
	Light	All	20,000	18,250	16,500	14,750	13,000
	Medium	All	24,000	23,038	22,075	21,113	20,150
	Heavy	All	40,000	36,906	33,813	30,719	27,625
CO_2 emission factor	All	Electricity (kg/kWh)	0.71	0.32	0.23	0.16	0.14

5. Results and Discussions

5.1. CO_2 Emissions of Subsectors of Road Transport in Different Scenarios

Based on the methods and data introduced above, CO_2 emissions from all passenger vehicles and freight vehicles in road transport can be calculated. For better comparison, we further divide passenger transport into subsectors of mini, small, medium, and large passenger vehicles and freight transport into subsectors of mini, light, medium, and heavy freight vehicles. Based on the LEAP, we can analyze CO_2 emissions from different subsectors of road transport in China from 2020 to 2060 under the eight policy scenarios proposed above. The results are shown in Figure 4.

Figure 4a shows that the total CO_2 emissions of road transport in the BAU scenario initially increase and then decrease yearly after reaching a peak of 1419.50 million tonnes in 2033. The CO_2 emissions of passenger transport increase and then decrease, from 54.08% in 2020 to 17.82% in 2060. On the other hand, the CO_2 emissions of freight transport grow steadily from 45.92% in 2020 to 82.18% in 2060. Before 2034, the CO_2 emissions of passenger transport are greater than those of freight transport. After 2034, the CO_2 emissions of passenger transport are gradually less than those of freight transport and continue to decline. The CO_2 emissions of small passenger vehicles and heavy freight vehicles are relatively large. While the CO_2 emissions of small passenger vehicles begin to decline after reaching a peak of 686.27 million tonnes in 2028, the growth trend of freight transport CO_2 emissions gradually slows down and begins to decline after 2056. The CO_2 emissions of other types of vehicles account for a relatively small proportion, with the trend relatively stable.

Figure 4b shows that in the CE scenario, due to the wide application of clean electricity for passenger vehicles, the reduction in passenger transport CO_2 emissions is significant. However, the CO_2 emission reduction of freight transport is small since the implementation of electric vehicles in freight transportation is more difficult. Figure 4c shows that the FEI scenario has promoted a great reduction in CO_2 emissions for all subsectors, with the CO_2 emission peak dropping significantly and the rate of CO_2 emission reduction accelerating after the peak. According to Figure 4d, in the SAV scenario, due to a decrease in the number of passenger vehicles and a change in vehicle structure, with an improvement in the fuel economy of vehicles, CO_2 emissions of passenger vehicles are significantly reduced compared with the BAU scenario. A comparison of Figure 4e–g shows that in the CET scenario, with an increase in carbon prices , the reduction of road transport CO_2 emissions becomes increasingly obvious. This is especially the case in the CET-HCP scenario, where the road transport CO_2 emissions will begin to decrease yearly after 2026 and will only be 488.10 million tonnes in 2060. Figure 4h clearly shows that in the CS scenario, under the combined effect of various emission reduction policies, the downward trend of CO_2 emissions from road transport is the quickest and the emission reduction is the most significant for passenger transport and freight transport.

5.2. CO_2 Emission Reduction Potential of Road Transport in Different Scenarios

5.2.1. Total CO_2 Emissions of Road Transport in Different Scenarios

Based on the LEAP, the total CO_2 emissions of China's road transport in different scenarios from 2020 to 2060 can also be analyzed and compared, as shown in Figure 5. The CO_2 emission gaps between the BAU scenario and other scenarios are then regarded as the CO_2 emission reduction potential.

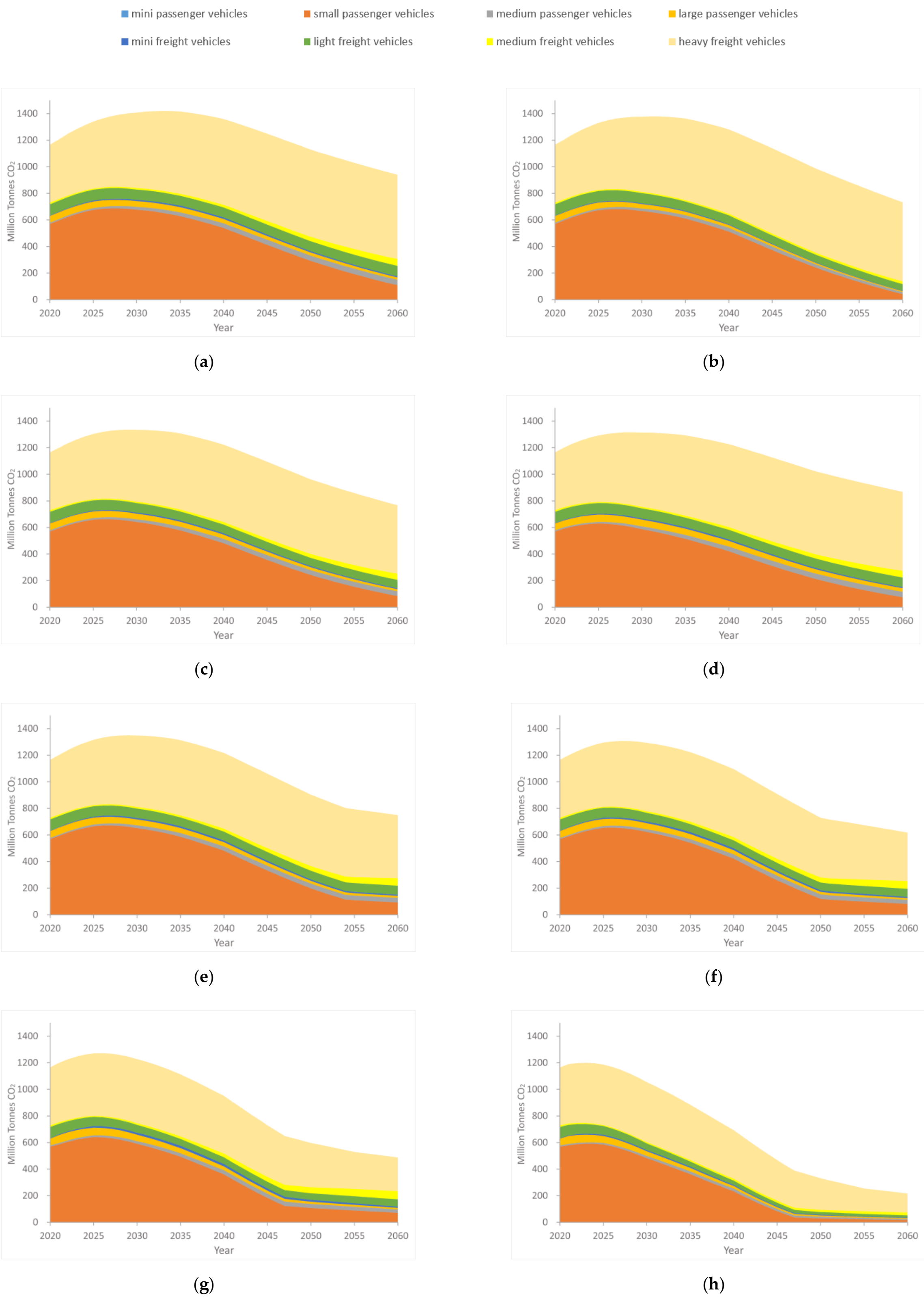


Figure 4. Projection of CO_2 emissions from subsectors of road transport in different scenarios. (**a**) Business-as-usual (BAU) scenario; (**b**) Clean electricity (CE) scenario; (**c**) Fuel economy improvement (FEI) scenario; (**d**) Shared autonomous vehicles (SAV) scenario; (**e**) Scenario of CET with low carbon prices (CET-LCP); (**f**) Scenario of CET with medium carbon prices (CET-MCP); (**g**) Scenario of CET with high carbon prices (CET-HCP); (**h**) Comprehensive (CS) scenario.

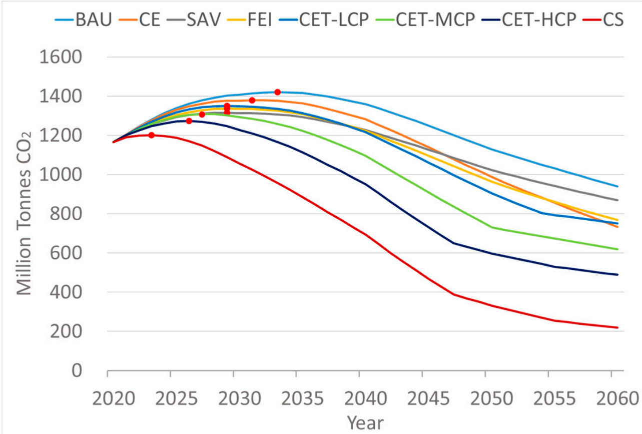


Figure 5. Projection of total CO_2 emissions from road transport in different scenarios.

Compared with the BAU scenario, the CO_2 emission reduction of road transport in the CE scenario is not significant before 2031 due to the relatively low penetration rate of electric vehicles. However, with the expansion of the vehicle electrification scale, the CO_2 emission of electric vehicles from power generation becomes one of the main emission sources in road transport. Therefore, the CO_2 emission reduction potential in the CE scenario is increasing yearly and even surpasses that in the FEI, SAV, and CET-LCP scenarios by 2060. The CO_2 emission reduction in the FEI scenario is relatively steady every year due to the continuing improvement in the fuel economy of vehicles, whereas the CO_2 emission reduction potential in the SAV scenario is gradually decreasing yearly since the vehicle population tends to be saturated. In the CET scenarios, the CO_2 emission reduction of road transport is gradually increasing. A comparison of CET-LCP, CET-MCP, and CET-HCP scenarios reveals that the CO_2 emission reduction potential increases with the rise of carbon prices. Among all the scenarios, the CS scenario has the greatest potential for road transport CO_2 emission reduction with all the policies implemented.

For better evaluation and comparison of the CO_2 emission reduction potential under different scenarios, the key numerical results are listed in Table 11. It shows that the peak years of CO_2 emission are becoming earlier from the BAU scenario to the CS scenario. The cumulative CO_2 emission in different scenarios are ranked as follows: BAU > CE > SAV > FE I > CET-LCP > CET-MLP > CET-HCP > CS. Among them, the CS scenario has the greatest potential for CO_2 emission reduction. In the CS scenario, CO_2 emissions of road transport will peak at 1200.37 million tonnes in 2023 and decrease to 217.73 million tonnes by 2060. The reduction rate of road transport CO_2 emission from the carbon peak year to 2060 can be up to 82%. From 2020 to 2060, the cumulative CO_2 emissions in the CS scenario are only 28,572.73 million tonnes. Compared with the BAU scenario, the cumulative CO_2 emission reduction is 22,501.22 million tonnes in the CS scenario.

Table 11. Comparison of road transport CO_2 emissions under different scenarios from 2020 to 2060.

Scenarios	Peak Year of CO_2 Emission	CO_2 Emission Peak (Million Tonnes)	Cumulative CO_2 Emission (Million Tonnes)	Cumulative CO_2 Emission Reductions Compared with the BAU Scenario (Million Tonnes)	CO_2 Emission Reduction Rate from Carbon Peak Year to 2060
BAU	2033	1419.50	51,073.75	-	34%
CE	2031	1378.69	47,460.70	3613.05	47%
SAV	2029	1315.47	47,274.21	3799.54	34%
FEI	2029	1335.51	46,320.29	4753.46	43%
CET-LCP	2029	1350.32	45,587.65	5486.10	44%
CET-MCP	2027	1307.91	41,560.01	9513.74	53%
CET-HCP	2026	1272.22	37,110.66	13,963.08	62%
CS	2023	1200.37	28,572.73	22,501.22	82%

5.2.2. Comparison with Previous Studies

In this study, we separately calculated CO_2 emissions of passenger transport and freight transport and obtained the total CO_2 emissions of road transport. In this section, we compare the results with some previous studies on the CO_2 emissions of passenger transport and freight transport in China. For example, Peng et al. [47] predicted that the direct CO_2 emissions of the road transport sector in mainland China will peak at 1500 million tonnes around 2030 and gradually decline to 1341.3 million tonnes in 2050 in the reference scenario. In the BAU and low carbon scenarios, the direct CO_2 emissions further decrease to 892.6 and 620.6 million tonnes in 2050, respectively. Gambhir et al. [48] demonstrated that road transport CO_2 emissions in China could decrease from 2080 million tonnes in the BAU scenario to 1240 million tonnes in the low carbon scenario by 2050. Yan et al. [49] indicated that CO_2 emissions of China's road transport sector in 2030 would reach 1303.7 million tonnes in the BAU scenario, which will be reduced to 783.1 million tonnes in the best-case scenario. Through the comparison, it can be found that the CO_2 emission reduction potential of China's road transport in our study is much greater than that in the previous studies. It indicates that the policy scenarios in our studies may have more significant effects on the CO_2 emission reduction of road transport.

5.3. Factor Contribution to Road Transport CO_2 Emission Reduction

Based on the LMDI method, the factors contributing to the emission reduction of road transport from the peak year to 2060 are composed for each scenario, as shown in Figures 6–13.

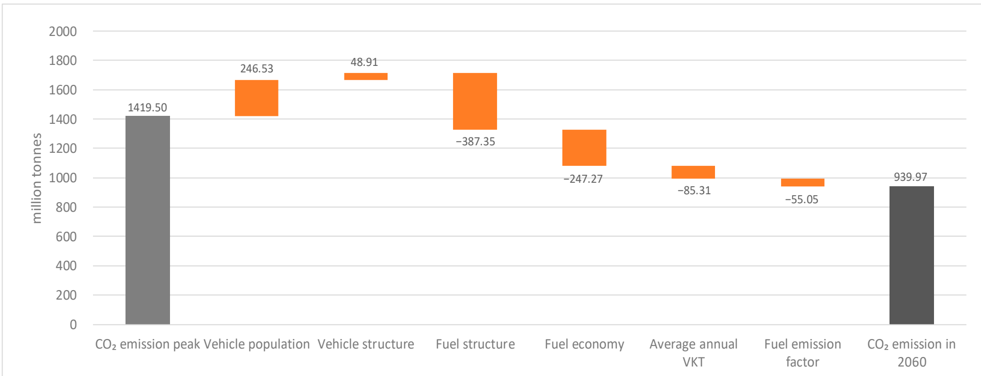


Figure 6. Factor contribution to road transport CO_2 emission reduction in the BAU scenario.

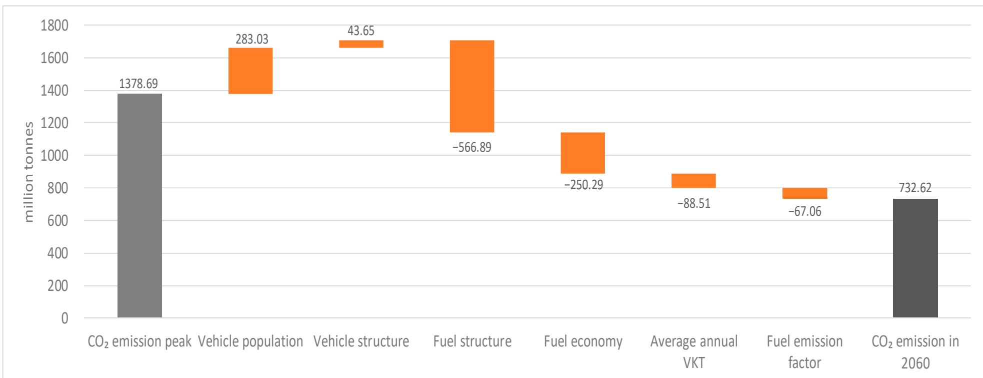


Figure 7. Factor contribution to road transport CO_2 emission reduction in the CE scenario.

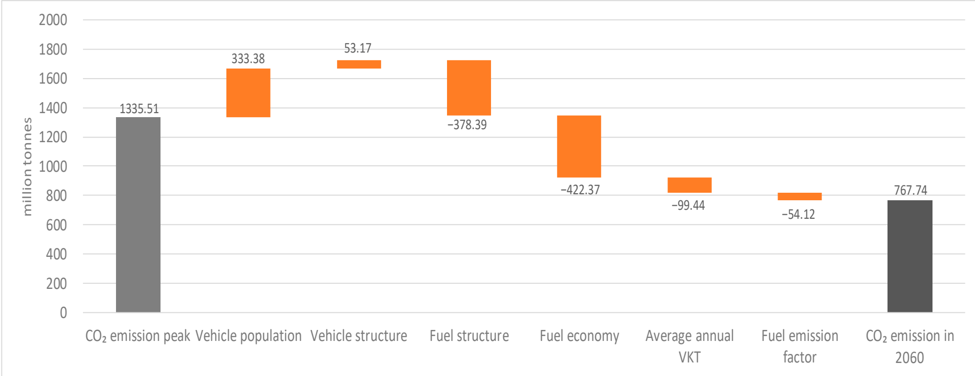


Figure 8. Factor contribution to road transport CO_2 emission reduction in the FEI scenario.

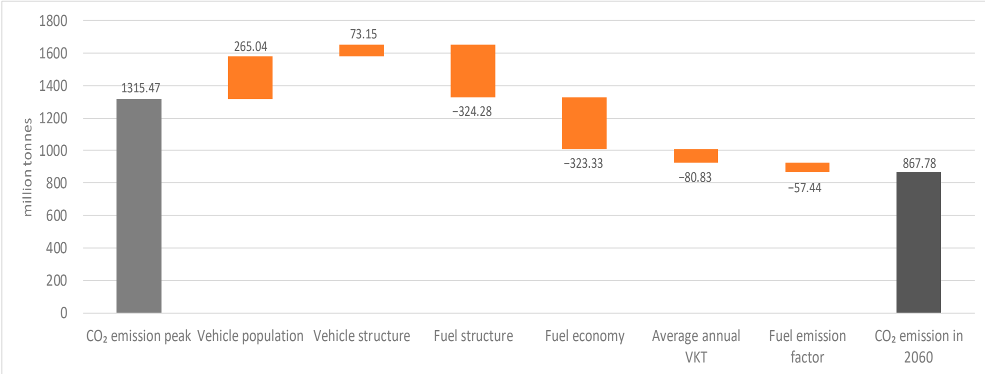


Figure 9. Factor contribution to road transport CO_2 emission reduction in the SAV scenario.

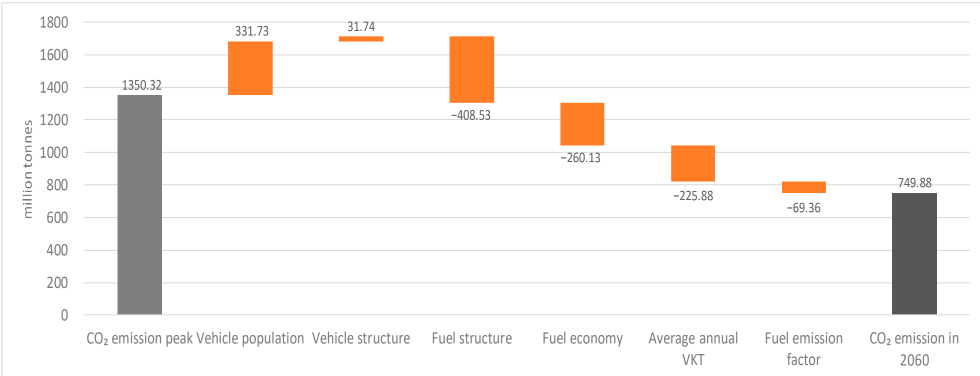


Figure 10. Factor contribution to road transport CO_2 emission reduction in the CET-LCP scenario.

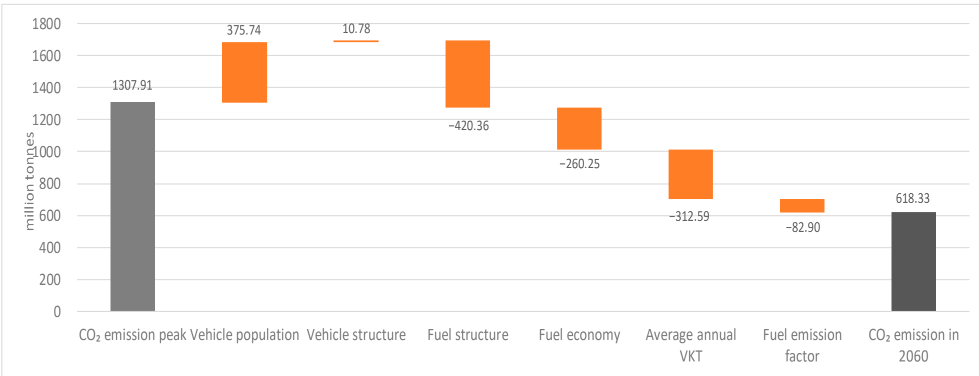


Figure 11. Factor contribution to road transport CO_2 emission reduction in the CET-MCP scenario.

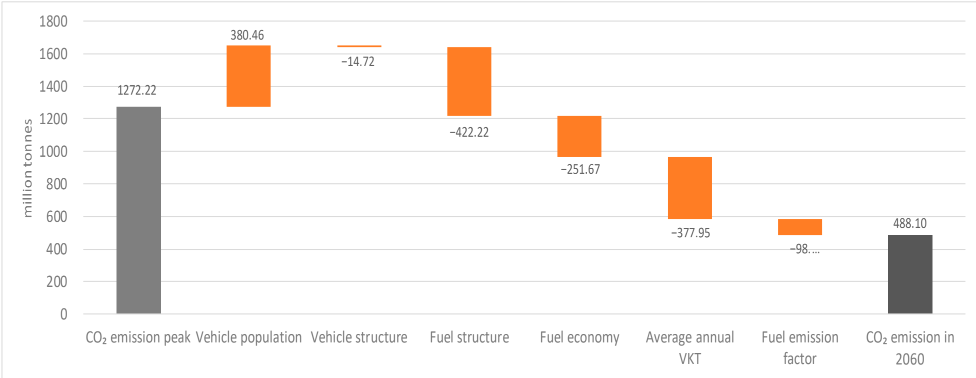


Figure 12. Factor contribution to road transport CO_2 emission reduction in the CET-HCP scenario.

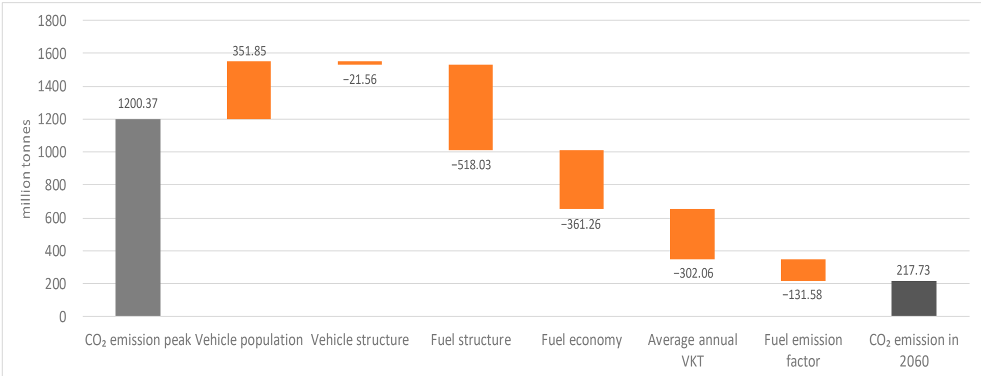


Figure 13. Factor contribution to road transport CO_2 emission reduction in the CS scenario.

As shown in Figure 6, among the various influencing factors of CO_2 emissions in the BAU scenario, the increase in vehicle population has a positive role in promoting an increase in CO_2 emissions, with the contribution of CO_2 emissions being as high as 246.53 million tonnes, which is the main reason for the increase in road transport CO_2 emissions. Besides, the vehicle structure also has a positive role in promoting an increase in CO_2 emissions. However, vehicle structure's contribution is only 48.91 million tonnes, substantially less than the influence of the vehicle population. The factors that contribute to the reduction of CO_2 emissions include the average annual VKT, fuel structure, fuel economy, and fuel emission factor. With an increase in the proportion of electric vehicles and other clean-energy vehicles, an improvement in fuel economy, and a decrease in the electricity CO_2 emission factor, road transport CO_2 emissions will be reduced. The most influential factor of CO_2 emission reduction is fuel structure, with its contribution reaching 387.35 million tonnes in the BAU scenario.

According to Figures 7–13, the contribution of factors to the road transport CO_2 emission reduction varies from scenario to scenario. In the CE scenario, the emission reduction contribution of the fuel emission factor increased by 12.01 million tonnes compared with the BAU scenario due to cleaner electricity used in the electric vehicles. Similarly, the contribution of fuel economy factors to emission reduction in the FEI scenario increased from 247.27 million tonnes (BAU) to 422.37 million tonnes since the additional improvement in the fuel economy of vehicles. The role of the vehicle population in the increase of road transport CO_2 emissions is weakened in the SAV scenario because of the reduced demand for private cars. The distributions of the factor contribution to road transport CO_2 emission reduction in the three CET scenarios are similar. Fuel structure is the driving factor of CO_2 emission reduction in CET scenarios with market-based incentives for low-carbon vehicles. In the CS scenario, the contribution of all factors to the road transport CO_2 emission reduction is enhanced.

To comparatively analyze the relative contribution of various factors, the contribution rates of each factor under different policy scenarios are shown in Figure 14. In all scenarios, the contribution rate of the vehicle population to CO_2 emissions is positive and greater than 0.35, which restrains the reduction of road transport CO_2 emissions. However, the average annual VKT, fuel structure, fuel economy, and fuel emission factor contribute to reducing road transport CO_2 emissions, which can promote the reduction of road transport CO_2 emissions. Fuel structure and fuel economy are the most important factors that restrain road transport CO_2 emissions. The contribution rate of fuel structure is up to −0.88 in the CE scenario, with the contribution rate of fuel economy up to −0.74 and −0.72 in the SAV and FE scenarios, respectively.

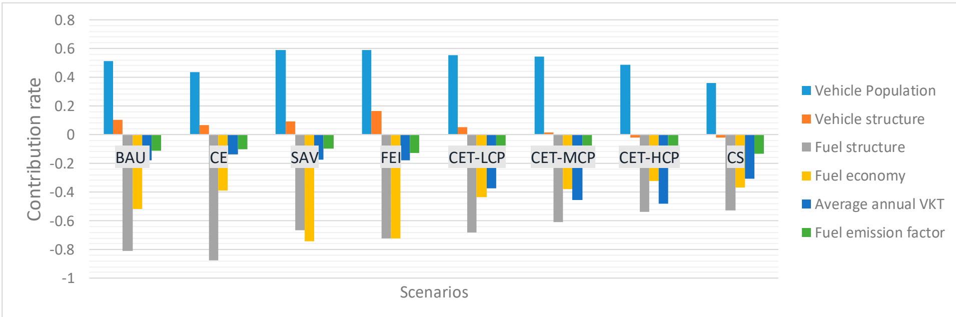


Figure 14. Contribution degree of factors to road transport CO_2 emissions.

To sum up, the vehicle population has the greatest impact on road transport CO_2 emission, followed by the vehicle structure, fuel economy, fuel structure, and average annual VKT. The emission reduction contribution of the fuel emission factor is relatively small, but with an increase in the proportion of electric vehicles, the fuel emission factor will be important and non-negligible in the medium and long term for the reduction of road transport CO_2 emission. Therefore, to achieve a deep and comprehensive reduction in road transport CO_2 emissions, all the factors discussed above should be considered when developing a low-carbon policy for road transport.

6. Conclusions and Policy Implications

To help achieve the goal of carbon neutrality in China, this study proposed eight policy scenarios to reduce CO_2 emissions of road transport. These scenarios are defined by several key factors that influence road transport CO_2 emissions, such as vehicle population, vehicle structure, fuel structure, fuel economy, average annual VKT, and fuel emission factors. Based on the scenarios set on the LEAP, road transport CO_2 emissions in China are analyzed from 2020 to 2060 for each scenario. Compared with the BAU scenario, the CO_2 emission reduction potential of the other seven scenarios is evaluated. Furthermore, the factors contributing to the emission reduction of road transport from the peak year to 2060 are composed using the LMDI method. The main findings are summarized as follows:

(1) Due to the widespread adoption of electric vehicles for passenger transport, they have a greater potential to reduce CO_2 emissions than freight transport in the field of road transport, especially for small passenger vehicles.

(2) The total CO_2 emissions of road transport will peak at 1419.5 million tonnes in 2033 for the BAU scenario. In contrast, the peaks of road transport CO_2 emissions for the CE, SAV, FEI, CET-LCP, CET-MCP, CET-HCP, and CS scenarios are decreasing and occur progressively earlier, as early as 2023.

(3) Compared with the BAU scenario, the cumulative CO_2 emission reductions of road transport from 2020–2060 for the other seven scenarios can be up to 22,501.22 million tonnes. The CO_2 emission reduction potential of the seven scenarios can be ranked as follows: CS > CET-HCP > CET-MCP > FEI > CET-LCP > SAV > CE. This finding

indicates that CO_2 emission trading may be more effective than other policies, with a combination of policies the best.

(4) Based on the decomposition of factors that contribute to the CO_2 emission reduction of road transport from the peak year to 2060 for each scenario, it is concluded that fuel structure and fuel economy contribute most to the emission reduction, whereas the increase in vehicle population restrains the CO_2 emission reduction.

The above findings also provide certain policy implications for the government to design a pathway toward low-carbon road transport under the goal of CO_2 emission peak and carbon neutrality in China. The specific policy suggestions are as follows:

(1) The power industry needs to vigorously increase the proportion of clean energy in power generation, including photovoltaic, hydroelectric, wind, and nuclear powers, to further reduce CO_2 emissions of electric vehicles.
(2) The government should formulate relevant policies to encourage vehicle manufacturers to improve the fuel economy of both traditional internal combustion engine vehicles and new energy vehicles to reduce the energy consumption and CO_2 emissions of road transportation.
(3) Since private vehicles account for a large proportion of passenger transport in China, the government could implement downstream emission trading for road transport to encourage more consumers to purchase new energy vehicles.
(4) To ensure the achievement of the targets of peak CO_2 emissions and carbon neutrality in China, a comprehensive policy package should be designed considering all the contributing factors to the emission reduction of road transport.

This study also has some limitations. Due to the immature technology and the high cost of hydrogen and other synthetic fuel cell vehicles, the future development of these vehicles is uncertain and difficult to predict. In addition, the proportion of these vehicles is currently negligible. Therefore, this study does not include hydrogen and other synthetic fuels in the fuel structure. The COVID-19 pandemic situation may reduce the traffic demand of road transport during the prevention and control period. However, the period was short and the impact was moderate across the year 2020. Therefore, this study does not consider the consequence of the COVID-19 pandemic in the scenario analysis of the road transport CO_2 emissions in China from 2020 to 2060, since we use the average annual VKT in the calculation. These limitations should be further addressed in future studies.

Author Contributions: Methodology, J.D. and W.L.; data curation, Y.L.; writing—original draft preparation, Y.L. and S.L.; writing—review and editing, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was sponsored by the National Natural Science Foundation of China (Grant No.: 52002244); Chenguang Program (Grant No.: 20CG55) supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission; and Shanghai Pujiang Program (Grant No.: 2020PJC083).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

C	total CO_2 emissions of road transport
C_P	total CO_2 emissions of road passenger transport
C_T	total CO_2 emissions of road freight transport
N_P	passenger vehicle population
S_i	proportion of passenger vehicles of type i
M_i	average annual VKT of type i vehicles

Q_{ik}	proportion of fuels of type k for vehicles of type i
F_{ik}	fuel economy of vehicles of type i with fuel type k
E_k	CO_2 emission factor of fuel of type k
N_T	freight vehicle population
S_j	proportion of freight vehicles of type j
M_j	average annual VKT of trucks of type j
Q_{jh}	proportion of fuels of type h for freight vehicles of type j
F_{jh}	fuel economy of trucks of type j with fuel type h
E_h	CO_2 emission factor of fuel of type h
ΔC	total contribution of each factor to road transport CO_2 emissions
ΔC_N	contribution of vehicle population to road transport CO_2 emissions
ΔC_S	contribution of vehicle structure to road transport CO_2 emissions
ΔC_Q	contribution of fuel structure to road transport CO_2 emissions
ΔC_F	contribution of fuel economy to road transport CO_2 emissions
ΔC_M	contribution of average annual VKT to road transport CO_2 emissions
ΔC_E	contribution of fuel CO_2 emission factors to road transport CO_2 emissions
i	type of passenger vehicles
j	type of freight vehicles
k	type of fuel used by passenger vehicles
h	type of fuel used by freight vehicles
BAU	business-as-usual scenario
CE	clean electricity scenario
FEI	fuel economy improvement scenario
SAV	shared autonomous vehicles Scenario
CET-LCP	CO_2 emissions trading scenario with low carbon prices
CET-MCP	CO_2 emissions trading scenario with mid carbon prices
CET-HCP	CO_2 emissions trading scenario with high carbon prices
CS	comprehensive scenario

References

1. European Automobile Manufacturers Association (ACEA). *Paving the Way to Carbon-Neutral Transport: 10-Point Plan to Help Implement the European Green Deal*; ACEA: Brussels, Belgium, 2020.
2. Li, W.; Bao, L.; Wang, L.; Li, Y.; Mai, X. Comparative evaluation of global low-carbon urban transport. *Technol. Forecast. Soc. Chang.* **2019**, *143*, 14–26. [CrossRef]
3. Li, W.; Bao, L.; Li, Y.; Si, H.; Li, Y.M. Assessing the transition to low-carbon urban transport: A global comparison. *Resour. Conserv. Recycl.* **2022**, *180*, 106179. [CrossRef]
4. Paustian, K.; Ravindranath, N.H.; Amstel, A.V. *2006 IPCC Guidelines for National Greenhouse Gas Inventories*; International Panel on Climate Change: Geneva, Switzerland, 2006.
5. Vuuren, D.; Hoogwijk, M.; Barker, T.; Riahi, K.; Boeters, S.; Chateau, J.; Scrieciu, S.; Vliet, J.; Masui, T.; Blok, K.; et al. Comparison of top-down and bottom-up estimates of sectoral and regional greenhouse gas emission reduction potentials. *Energy Policy* **2009**, *37*, 5125–5139. [CrossRef]
6. Li, Y.; Bao, L.; Li, W.; Deng, H. Inventory and policy reduction potential of greenhouse gas and pollutant emissions of road transportation industry in China. *Sustainability* **2016**, *8*, 1218. [CrossRef]
7. Shan, Y.; Guan, D.; Liu, J.; Mi, Z.; Liu, Z.; Liu, J.; Schroeder, H.; Cai, B.; Chen, Y.; Shao, S.; et al. Methodology and applications of city level CO_2 emission accounts in China. *J. Clean. Prod.* **2017**, *161*, 1215–1225. [CrossRef]
8. Brand, C.; Tran, M.; Anable, J. The UK transport carbon model: An integrated life cycle approach to explore low carbon futures. *Energy Policy* **2012**, *41*, 107–124. [CrossRef]
9. Zhang, L.; Long, R.; Chen, H.; Geng, J. A review of China's road traffic carbon emissions. *J. Clean. Prod.* **2019**, *207*, 569–581. [CrossRef]
10. Liu, M.; Zhang, X.; Zhang, M.; Feng, Y.; Liu, Y.; Wen, J.; Liu, L. Influencing factors of carbon emissions in transportation industry based on CD function and LMDI decomposition model: China as an example. *Environ. Impact Assess. Rev.* **2021**, *90*, 106623. [CrossRef]
11. Quan, C.; Cheng, X.; Yu, S.; Ye, X. Analysis on the influencing factors of carbon emission in China's logistics industry based on LMDI method. *Sci. Total Environ.* **2020**, *734*, 138473. [CrossRef]
12. Brand, C.; Anable, J.; Tran, M. Accelerating the transformation to a low carbon passenger transport system: The role of car purchase taxes, feebates, road taxes and scrappage incentives in the UK. *Transp. Res. Part A Policy Pract.* **2013**, *49*, 132–148. [CrossRef]

13. Dineen, D.; Ryan, L.; Gallachóir, B.Ó. Vehicle tax policies and new passenger car CO_2 performance in EU member states. *Clim. Policy* **2018**, *18*, 396–412. [CrossRef]
14. Nayum, A.; Klöckner, C.A.; Prugsamatz, S. Influences of car type class and carbon dioxide emission levels on purchases of new cars: A retrospective analysis of car purchases in Norway. *Transp. Res. Part A Policy Pract.* **2013**, *48*, 96–108. [CrossRef]
15. Patil, G.R. Emission-based static traffic assignment models. *Environ. Model. Assess.* **2016**, *21*, 629–642. [CrossRef]
16. Roman, R.D.; Ritchie, S.G. Accounting for population exposure to vehicle-generated pollutants and environmental equity in the toll design problem. *Int. J. Sustain. Transp.* **2017**, *11*, 406–421. [CrossRef]
17. Yan, S.; Eskeland, G.S. Greening the vehicle fleet: Norway's CO_2-Differentiated registration tax. *J. Environ. Econ. Manag.* **2018**, *91*, 247–262. [CrossRef]
18. Hörcher, D.; Borger, B.D.; Seifu, W.; Graham, D.J. Public transport provision under agglomeration economies. *Reg. Sci. Urban Econ.* **2020**, *81*, 103503. [CrossRef]
19. Timilsina, G.R.; Dulal, H.B. *Fiscal Policy Instruments for Reducing Congestion and Atmospheric Emissions in the Transport Sector: A Review*; World Bank Policy Research Working Paper No. 4652; World Bank: Washington, DC, USA, 2010; Available online: https://ssrn.com/abstract=1149604 (accessed on 30 April 2022).
20. Chen, S.; Zheng, X.; Yin, H.; Liu, Y. Did Chinese cities that implemented driving restrictions see reductions in PM10? *Transp. Res. Part D Transp. Environ.* **2020**, *79*, 102208. [CrossRef]
21. Burniaux, J.M.; Chateau, J. *Mitigation Potential of Removing Fossil Fuel Subsidies: A General Equilibrium Assessment*; OECD Economics Department Working Papers; OECD: Paris, France, 2011.
22. Jia, S.; Mao, B.; Liu, S.; Sun, Q. Calculation and analysis of transportation energy consumption level in China. *J. Transp. Syst. Eng. Inf. Technol.* **2010**, *10*, 22–27. [CrossRef]
23. Li, Y.; Bao, L.; Bao, J. Evaluating Emission Mitigation Potential of Shanghai Transportation Policies Using LEAP. In Proceedings of the 93rd Transportation Research Board, Washington, DC, USA, 12–16 January 2014.
24. Liu, H.; Wang, Y.; Chen, X.; Han, S. Vehicle emission and near-road air quality modeling in Shanghai, China, based on taxi GPS data and MOVES revised emission inventory. *Transp. Res. Rec.* **2013**, *2340*, 38–48. [CrossRef]
25. Feng, X.; Zhao, M.; Wang, M.; Du, X.; Tian, C.; Gao, J. Simulation research on co-controlling pollutants and greenhouse gases emission in China's transportation sector. *Clim. Change Res.* **2021**, *17*, 279–288.
26. Pang, K.; Zhang, Q.; Ma, C.; Zhu, L.; Chen, H.; Kong, X.; Pan, F.; Yang, H. Forecasting of Emission Co-reduction of Greenhouse Gases and Pollutants for the Road Transport Sector in Lanzhou Based on the LEAP Model. *Environ. Sci.* **2021**. (In Chinese) [CrossRef]
27. Duan, L.; Hu, W.; Deng, D.; Fang, W.; Xiong, M.; Lu, P.; Li, Z.; Zhai, C. Impacts of reducing air pollutants and CO_2 emissions in urban road transport through 2035 in Chongqing, China. *Environ. Sci. Ecotechnol.* **2021**, *8*, 100125. [CrossRef]
28. Chi, L. The Prediction of Beijing Passenger Transport Energy Demand and Pollutants Emission in Future Based on LEAP Mode. Ph.D. Thesis, Beijing Jiaotong University, Beijing, China, 2014. (In Chinese).
29. Hu, X.L.; Jiang, K.J. CO_2 emission reduction strategies analysis: AIM/energy-emission model. *Energy China* **1998**, *11*, 17–22.
30. Zhang, C.X. Research on the Development of Electric Vehicle Impact on Energy and Environment in Guangzhou Based on LEAP. Master's Thesis, North China Electric Power University, Beijing, China, 2013. (In Chinese).
31. Zhang, M.; Miu, H. Research on the Change of Motorized Passenger Transport Energy Consumption in Selected Chinese Cities Based on LMDI Method. *Acta Sci. Nat. Univ. Pekin.* **2010**, *46*, 483–486. [CrossRef]
32. International Energy Agency (IEA). *The Future of Trucks: Implications for Energy and the Environment*; IEA: Paris, France, 2017.
33. Pierpaolo, C.; Marine, G.; Jacopo, T. *Global EV Outlook 2019—Scaling Up the Transition to Electric Mobility*; IEA: Paris, France, 2019.
34. International Energy Agency (IEA). *Global Energy and CO_2 Status Report 2018: The Latest Trends in Energy and Emissions in 2018*; IEA: Paris, France, 2019.
35. Li, R.; He, Q.; Li, S. Analyses of Automobile Ownership Development Trend in China. *Urban Transp. China* **2013**, *11*, 69–75. [CrossRef]
36. China International Capital Corporation Limited. Passenger Car Ownership is Expected to Rise and then Fall to 530 Million in 2060. Available online: http://stock.jrj.com.cn/2020/12/04081131405001.shtml (accessed on 29 January 2022).
37. An, F.; Kang, L.; Qin, Z.; Mao, S.; Wang, W.; Dror, M.B. Study on the phasing-out timetable of China's traditional ICE-vehicles. *Int. Pet. Econ.* **2019**, *27*, 1–8.
38. Updated Timetable for Phasing Out China's Traditional Ice Vehicles and Its Environmental Benefits Assessment. Available online: http://www.nrdc.cn/Public/uploads/2020-06-30/5efac9b5b0832.pdf (accessed on 29 January 2022).
39. Ministry of Industry and Information Technology of the People's Republic of China. *Fuel Consumption Limits for Passenger Cars GB19578-2021*; Ministry of Industry and Information Technology: Beijing, China, 2021.
40. Toward "Net Zero" Emissions in the Road Transport Sector in China. Available online: https://wri.org.cn/research/toward-net-zero-emissions-road-transport-sector-china (accessed on 29 January 2022).
41. Jia, S.; Song, Y.; Yang, Z. Process of Peak carbon emissions of trucks during operating activities based on growth curve function. *J. Transp. Syst. Eng. Inf. Technol.* **2021**, *21*, 310. [CrossRef]
42. Li, W.; Pu, Z.; Li, Y.; Ban, X. Characterization of ridesplitting based on observed data: A case study of Chengdu, China. *Transp. Res. Part C Emerg. Technol.* **2019**, *100*, 330–353. [CrossRef]

43. Li, W.; Pu, Z.; Li, Y.; Tu, M. How does ridesplitting reduce emissions from ridesourcing? A spatiotemporal analysis in Chengdu, China. *Transp. Res. Part D Transp. Environ.* **2021**, *96*, 102885. [CrossRef]
44. Li, W.; Chen, S.; Dong, J.; Wu, J. Exploring the spatial variations of transfer distances between dockless bike-sharing systems and metros. *J. Transp. Geogr.* **2021**, *92*, 103032. [CrossRef]
45. Li, W.; Wang, L.; Li, Y.; Liu, B. A Blockchain-based emissions trading system for the road transport sector: Policy design and evaluation. *Clim. Policy* **2021**, *21*, 337–352. [CrossRef]
46. Song, L.; Cao, Z.; Dai, M. Material metabolism and carbon emission reduction strategies of passenger cars in China's mainland. *Resour. Sci.* **2021**, *43*, 501–512. [CrossRef]
47. Peng, T.; Ou, X.; Yuan, Z.; Yan, X.; Zhang, X. Development and application of China provincial road transport energy demand and GHG emissions analysis model. *Appl. Energy* **2018**, *222*, 313–328. [CrossRef]
48. Gambhir, A.; Tse, L.; Tong, D.; Martinez-Botas, R. Reducing China's road transport sector CO_2 emissions to 2050: Technologies, costs and decomposition analysis. *Appl. Energy* **2015**, *157*, 905–917. [CrossRef]
49. Yan, X.; Crookes, R.J. Reduction potentials of energy demand and GHG emissions in China's road transport sector. *Energy Policy* **2019**, *37*, 658–668. [CrossRef]

sustainability

Article

Nonlinear Rail Accessibility and Road Spatial Pattern Effects on House Prices

Kaida Chen [1,2], Hanliang Lin [1], Lingyun Liao [2,*], Yichen Lu [2], Yen-Jong Chen [1], Zehua Lin [2], Linxi Teng [2], Aifang Weng [2] and Tianqi Fu [2]

1 Department of Urban Planning, National Cheng Kung University, Tainan 70101, Taiwan; kaidachenwhu@gmail.com (K.C.); linx@mail.ncku.edu.tw (H.L.); yj_chen@mail.ncku.edu.tw (Y.-J.C.)
2 College of Arts &College of Landscape Architecture, Fujian Agriculture and Forestry University, Fuzhou 350002, China; lucyy688@sina.com (Y.L.); 18860153366@sina.cn (Z.L.); tcl0610@163.com (L.T.); 3191726046@fafu.edu.cn (A.W.); 1201775009@fafu.edu.cn (T.F.)
* Correspondence: liaolingyun@fafu.edu.cn

Abstract: The continuous change process in the impact of differences in public transport accessibility has not been explained specifically in previous studies. This study reveals that the interaction between two continuous explanatory variables has a significant impact on the explained variable in the hedonic model. The study takes the accessibility variable in the house price model as an instance, dividing the accessibility variable of the residential community into two parts. The first part is the rail accessibility defined by the Euclidean distance from the residential community to the nearest rail transportation station. The second part is the road accessibility defined by two Space Syntax indicators, connectivity and carrying capacity, according to the spatial pattern of the road network. As demonstrated by the spatial interactive regression model, this research finds that road connectivity has a significant regulating effect on the impact of the distance to the closest rail station on house prices based on the empirical evidence from Fuzhou, China.

Keywords: house price; transportation accessibility; space syntax; hedonic models; spatial interactive regression model

check for updates

Citation: Chen, K.; Lin, H.; Liao, L.; Lu, Y.; Chen, Y.-J.; Lin, Z.; Teng, L.; Weng, A.; Fu, T. Nonlinear Rail Accessibility and Road Spatial Pattern Effects on House Prices. *Sustainability* **2022**, *14*, 4700. https://doi.org/10.3390/su14084700

Academic Editors: Linchuan Yang, Yuanyuan Guo, Yaoming Zhou, Wenxiang Li and Jixiang Liu

Received: 8 March 2022
Accepted: 11 April 2022
Published: 14 April 2022

1. Introduction

The discussion of housing is an unavoidable topic of human development. Among the many topics in the discussion of housing, exploring the direction of house prices has always been a key component. In order to allow people more reasonable living costs, and under the premise of sustainable income control, finding suitable living conditions for their own needs has become an important reason for ordinary people to discuss house prices [1]. The government also pays attention to the issue of housing prices in order to balance the interests of different parts of the city through the location selection of public facilities, so as to maintain social harmony in a long-term sustainable manner [2]. Simultaneously, researchers hope to use house prices as a reference standard to measure the value of space so as to conduct long-term exploration of variables that affect it, which allows them to maintain a high enthusiasm for the issue of property prices [3]. Consequently, in essence, the reason why different groups pay attention to property prices is to maintain a high-quality living environment in a long-term sustainable manner.

There are many variables that affect housing prices. Whether it is a variable factor that has been confirmed or one that has not, the research will continue to be refined and expanded over time. Furthermore, different variables have not only a single linear relationship to the impact of house prices, but also due to overlapping or mutually exclusive functions between different variables so that there is a moderating effect between them. Thus, it is very important to explore the variables that affect house prices and the interactive moderating effects between variables.

Due to the accessibility attribute having a significant impact on the value of the property, the impact of accessibility variables on house prices has been a hot research direction for a long time [4]. Among research documenting in the impact of accessibility variables on house prices, rail accessibility and road accessibility are two main factors which define accessibility. Therefore, this study will take rail accessibility and road accessibility as examples to reveal the mutual moderating relationship between the two variables on the impact of house prices.

Scholars explore how property prices are affected by these accessibility variables respectively from familiar angles. These angles often need to be able to specifically reflect the characteristics and accuracy of residential objects. When analyzing impact of subway accessibility variables on property price, most previous studies reveal the influence scope and extent of subway accessibility variables on property prices by calculating the Euclidean distance between rail stations or road travel distance. To be specific, Ronghui Tan et al., (2019) researched subway traffic in Wuhan, finding that ordinary rail stations affected the price of surrounding properties within 1600 m, while transfer stations influenced that within 4800 m [5]. Moreover, Elif Can Cengiz et al., (2019) pointed out that the price of properties was higher near subway station facilities than those in the distance [6]. Since the road's spatial pattern is very decisive for road accessibility, the quantitative evaluation of the spatial pattern form can deeply reflect the degree of road accessibility. Thus, in terms of research on the impact of road accessibility on property prices, researchers focus on the influence of characteristics of road network forms. In this regard, Law (2017) investigated London and drew a conclusion that property price was significantly affected by the perception of difference in road network forms, and the influence became stronger over time [7]. Another example is Liang, who proved through empirical evidence in Victoria, Australia that the removal of urban level crossings changed traffic connectivity and road accessibility, resulting in an impact on housing prices [8]. However, the increase in traffic connectivity and road accessibility emphasized in Liang's research is a local judgment based on human intuition, rather than a geometric evaluation based on the overall urban road network. Therefore, the study will be carried out based on the explanatory variables with reference to the rail accessibility reflected by the Euclidean distance between the residence and the rail station and the road accessibility reflected by the road spatial pattern of the road network where the residence is located.

Among the existing literature, when discussing the impact of two or more explanatory variables on house prices due to their interactions, scholars mainly compare the difference in interactions by building a regression model in groups or adding dummy variable items. However, it is impossible to accurately assess regulatory effects of continuous variables. Specifically, Jen-Jia Lin et al. modeled data by groups of properties along Metro Taipei Redline before the opening (1993–1995) and after the opening (1997–1999), in order to reveal the impact of the Taipei Rapid Transit System on house prices before and after the opening [9]. Although the impact was different, the continuous change process of this difference could not be explained.

In general, it is of great value to adopt interaction items in econometrics to analyze how explained variables are affected by interaction of multiple non-colinear continuous explanatory variables. As mentioned in the beginning, urban transportation mainly includes subway and roads whose construction helped to improve the carrying capacity of the urban transportation system and drive the regional economy [10]. At the same time, houses or properties are a key component of urban areas. Therefore, it is significant to study urban development via revealing the impact of subway accessibility and road accessibility on house prices by the interaction principle. As for urban residents, they pay great attention to availability when purchasing houses. They want to know the direct impact of subway accessibility and road accessibility, but are eager to figure out the regulatory effect of interactions on house prices, so as to select the best house locations in accordance with economic strength and transportation demand preferences.

Therefore, this study introduced the concept of interaction into the hedonic price model, and established linear interactive regression models and a spatial interactive regression model for analyzing non-collinear continuous explanatory variables, so as to explain the mutual regulatory effects of the subway traffic accessibility variable and road accessibility variables on house prices. It is hoped that this will provide a valuable reference for government transportation construction, buyers purchasing properties and literature references of scholars. Due to the construction of urban transport, there is a significant increase in the value of urban land [11]. Therefore, by studying the differences in land value enhancement, it is possible to help the government to issue targeted land fiscal policies [12]. These land finance policies can help the government to obtain effective financing compensation for transportation construction, so as to continue to promote urban development and balance the development level of the region [13]. For individual home buyers, in addition to directly obtaining the reference for their own living choices, they can also obtain advice for their own investment in real estate, so as to obtain the maximum land investment income. For scholars, capturing these differences in land value improvement can better improve the theoretical system of spatial economics and establish more diverse disciplines, such as economics, planning, geography, sociology and so on.

The remaining research content falls into four sections. Section 2 presents a literature review for the research basis. Section 3 discusses the research framework, data sources and process. In Section 4 the research results obtained are presented. Section 5 provides discussions of the findings and of limitations based on the data results, and concludes with the implications of the study.

2. Literature Review

After the introduction of the research background, this study will review the literature discussing the application of the hedonic price model to study influences on house prices, interactive models of house prices in the past, the impact of the distance of subway stations as well as road network forms.

It is a very common method to study house prices through hedonic price models. Hedonic price model is a research method that decomposes the value composition of an item, and constitutes an item model by establishing the values of multiple features to evaluate the total price of the item. It was composed of consumer theory proposed by American scholar Lancaster in 1966 and the supply–demand equilibrium model by American economist Rosen. The earliest hedonic price model was often used to analyze price composition of commodities such as vegetables and automobiles [14]. It was not widely applied in the analysis and research of house prices until Ridker et al. [15] first studied the impact of the improvement of environmental quality on house prices through the model. Most researchers divided it into characteristic variables, a neighborhood characteristic variable and a location accessibility characteristic variable [16]. As a recent example, Cui et al. [17] used the hedonic price model to demonstrate that different types of consumers have different demand preferences for housing. Zou et al. [18] demonstrate the impact of environmental quality on housing rents. Claudia Hitaja et al. [19] and Wang and Lee [20] through their studies on cities in the United States and China, respectively, proved that air quality levels also have a real impact on housing prices. Sumit Agarwal et al. [21] in Singapore confirmed the impact of school districts on housing prices. In addition, the hedonic price model will also be applied to the analysis of other living environments, which will also indirectly affect the regional impact on housing prices. For instance, Sumit Agarwal and Koo Kang Mo. [22] analyzed periodic congestion charging in the context of Singapore, which concerned the impact of rate adjustments on changes in commuters' mode of transportation choices.

In the process of researching housing prices by previous scholars, interaction models have been used to study housing prices. In the traditional general hedonic price model, it is mainly used to investigate the influence of non-collinear explanatory variables on explained variables, while the interactive regression model serves for the influence of correlation of multiple variables with explained variables. However, in the past, the interaction item model was frequently adopted to explore the impact of interaction between non-continuous dummy variables on explained variables, while an interactive regression model defined by continuous variables was relatively rare. To be specific, Ting Xu [23] introduced an interactive model of dummy variables defined by environmental location and household classification, to prove that the marginal price of housing attributes in Shenzhen varied with the absolute location environment and household condition rather than being invariable. Similarly, Janet Currie et al. [24] defined the opening status of factories and explored whether properties were affected by the factory through dummy variables. It was found that prices of properties within 0.5 km of factories emitting toxic gases declined by appropriately 11%. Meanwhile, concerning the distance of properties from new subway stations and house transaction time, Junhong Im and Sung Hyo Hong [25] grouped property transaction data in Daegu, South Korea, to compare the premium of house prices by new subway stations under the influence of old subway stations. They classified the distance between properties and the new subway station and time difference before and after the announcement of new subway station policy into dummy variables, and then established an interactive regression model. Finally, a conclusion was drawn that a residence price premium was obvious for properties within 500 m of new subway stations and 5000 m away from original subway stations.

The impact of rail stations on house prices has been a hot research topic in recent decades. As a major form of infrastructure, urban subway stations have produced a significant impact on firm location, employment and population growth. This has been demonstrated by Mayer, T. and Trevien, C. [26] through empirical evidence in Paris. Similarly, the impact on house prices is also quite significant. Regarding this, Zheng Jiefen and Liu Hongyu [27] examined and analyzed data of Shenzhen subway, realizing that prices of properties within a 400m to 600m radius of subway stations were higher. With the change in distance from subway stations, price rising was dramatically different. Xinyue Zhang and Jiang Yanqing [28] researched data of second-hand houses within 6 km of Nanjing Metro Line 1 and Line 2, and verified that when the distance was less than 500 m, house price rose highest, and rising was still obvious within 2 km. With distances greater than 2 km, this influence was negligible. Similarly, Almosaind et al. [29] analyzed house price data around Portland's subway transit stations to find that price of properties within 500 m of subway transit stations soared by 10.6%. Haizhen Wen et al. [30] studied house price data along Hangzhou subway stations and concluded that house prices were negatively correlated with the distance to subway stations; that is, for every 1% decrease in distance, prices would rise by 0.049%. As the distance increased, the value-added effect was weaker. More precisely, properties within 500 m had the highest premium (6.2%). While the distance was over 1500 m, the value-added rate dropped sharply (2.1%). Armstrong and Rodriguez [31], through a study of 1860 residential samples in Massachusetts, USA, proved that the premium capacity of rail stations for residences is not the same outside the half-mile buffer zone of rail stations. In addition, differences in distance and degree of house prices affected by subway stations would also change with variation in other external conditions. Having analyzed the transaction prices of 722 residential communities near Shenzhen subway stations, Yang et al. [32] discovered that the price of surrounding properties was positively affected by rail accessibility (1200 m in urban area and 1600 m in the suburbs). The influence became weaker with properties more distant from the urban subway station.

Previous scholars have also tried to study the impact of road network morphology on house prices. Traditional road traffic accounts for the largest proportion in urban traffic systems. It is one major way for people to travel, so it affects house prices greatly. Many scholars have conducted studies in this regard. However, due to the display complexity of roads on a three-dimensional level, they are usually quantized from a two-dimensional plan. Relevant research began earlier in Western countries; for instance, Matthews et al. [33] conducted an empirical study of Seattle, finding that house price was affected by the road network form variable when at a walking scale of about 500 m. Xiao et al. [34]

concluded that improvement of road network accessibility improved prices. Moreover, taking Shanghai as a case, Yao Shen and Kayvan Karimi [35] explained how house prices in different areas were influenced by road network centrality and functional connectivity based on traditional spatial syntactic analysis. In China, the research on relationship between road network form and house price was still in its infancy. The hedonic price model was introduced by Xiao Yang et al. [36] to study Nanjing, and a conclusion was drawn that road network structure produced two economic effects on house prices in the downtown area. In addition, Yang Ying et al. [37] analyzed geographic weighted regression (GWR) and realized that the house price of second-hand houses was significantly affected by road network forms in Chengnan and Chengbei Districts of Xi'an City. Based on the sDNA model, Gu Hengyu et al. [38] discussed the influences of road network forms, and summarized that house price was impacted differently due to different description methods of road network forms.

3. Methods

This study was based on the idea of "hypothesis validation". Through the collected data, a spatial regression interaction model was established to explore the significance of the interaction terms of the target variables, so as to give relevant research results. Finally, the study ensured the stability of the results through Robustness checks.

3.1. Research Framework

According to the literature review, the study proposed the research hypothesis that "road network forms play different roles in regulating the extent of the influence of distance to subway stations on house prices, and vice versa". Based on the hypothesis, this research process includes five steps (Figure 1). The first step is road network spatial data reflecting the realities of roads, and the spatial syntax analysis method is applied to quantitatively collect urban road accessibility and other variables impacting house prices. In the second step, on the basis of the collected explanatory variables of properties, a hedonic price regression model is applied; interaction of accessibility variables was used to construct a linear regression interaction model for house prices. According to the linear interactive regression model, the third step deduces generality of space, and considers spatial auto-correlation of explained variable (house price) and spatial nuisance of error to build a spatial interactive regression model. The fourth step was to perform a robustness test of the spatial regression interaction model. The last step was to analyze and summarize based on the model regression coefficients obtained above.

The above research process can be applied to the variable interaction research of the hedonic price model of house price, no matter whether the variable is a continuous variable or a dummy variable. This is a further attempt compared to previous scholars who only used dummy variables to establish interaction models when they introduced interaction variables to discuss the impact of house prices.

On the other hand, the research process also has some limitations. One of the limitations of this research process is that the third step model of the research is based on spatial auto-correlation theory, which uses cross-sectional spatial data for spatial analysis, which fails to take temporal variables into account. In addition, there are other specific limitations of the study, which will be discussed in detail in the final chapter of the paper.

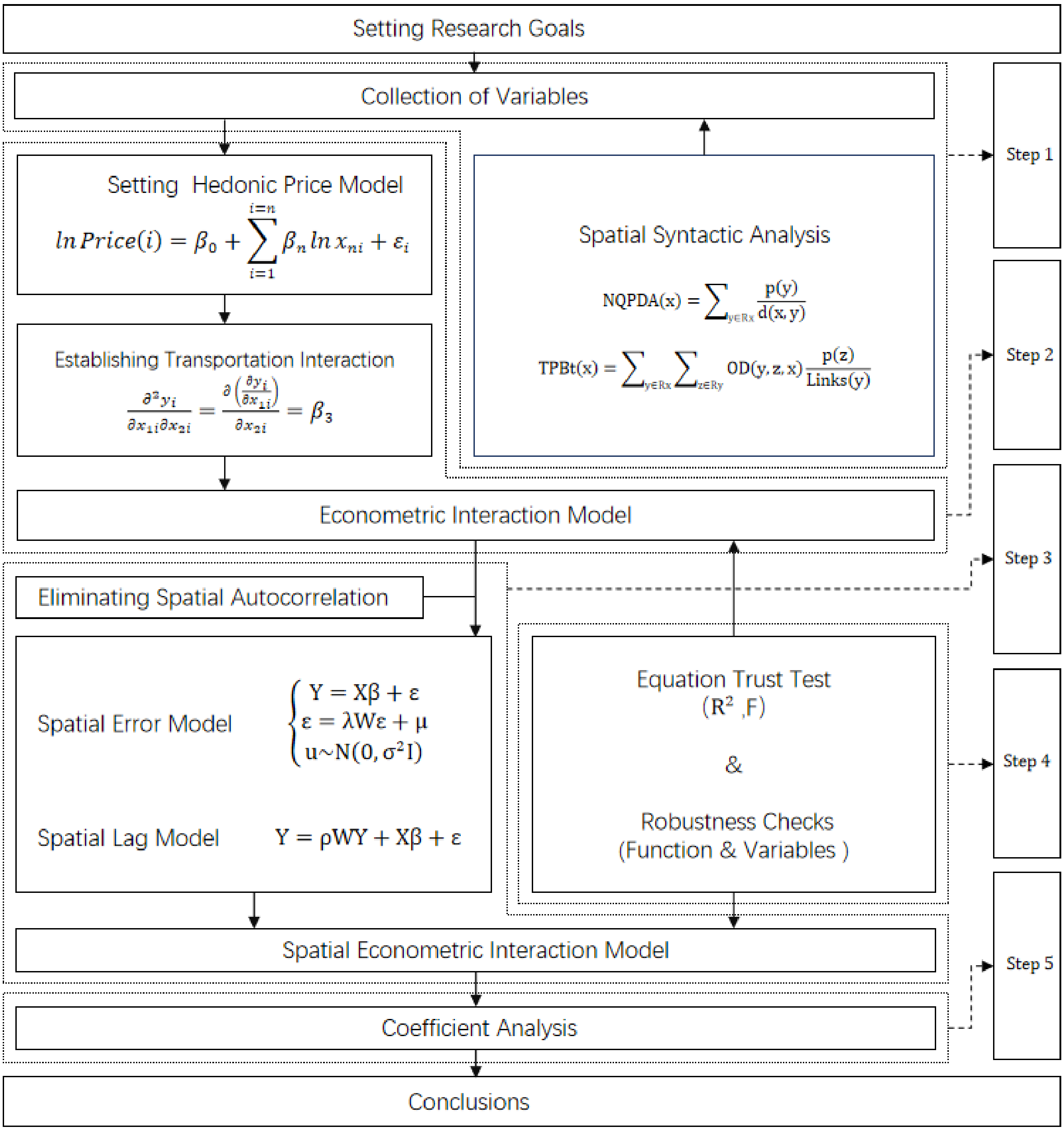


Figure 1. Frame diagram.

3.2. *Methodology*

3.2.1. Space Syntax Analysis

After establishing the research framework, this study adopted space syntax to calculate two sets of data of road accessibility—closeness, reflecting road connectivity, and betweenness, representing road carrying capacity [39]. According to space syntax theory, the definition of space was understood from scale and space levels, and it helped to further analyze the complex relationship between scale and space. In this study, mainly road space was analyzed.

Calculation of closeness means the efficiency and convenience of a certain road segment to search for other roads to be reached within search radius R. In the study, the value is represented by NQPDA (Network Quantity Penalized for Distance Analysis). The global closeness (NQPDAn, abbreviated as NQ) is that search radius R value is closeness value of the whole research scope. NQ shall be used to indicate the first measurement indicator of road segment in road accessibility (connectivity). In order to make the expression process simple, this study abbreviated the concept of NQ as closeness, with a calculation parameter model as follows:

$$\mathrm{NQPDA}(\mathrm{x}) = \sum\nolimits_{\mathrm{y}\in\mathrm{Rx}} \frac{\mathrm{p}(\mathrm{y})}{\mathrm{d}(\mathrm{x},\mathrm{y})} \quad (1)$$

wherein, NQPDA(x) represents closeness; p(y) means relative importance of index of node y in overall data evaluation within the range of search radius R; d(x,y) means the shortest topological path between node x and node y; Rx is the area with x as the center and R as radius.

Betweenness can be understood as the passing probability of traffic passing from one point to another of the road section within search radius R; it will be represented by TPBt (abbreviated to TP, the sum of geodesics that pass through the link from one point to another). Therefore, in the spatial syntax analysis model, the higher the value of betweenness, the greater the possible traffic volume carried by the road, and vice versa. In the spatial syntax line segment model, betweenness describes road carrying capacity in urban traffic, and is used in analysis of traffic flow and traffic evacuation capacity. This is the second measurement indicator to describe road accessibility. In addition, for easy expression, the global travel degree used in the study is referred to as betweenness in the following process, with calculation parameter model as below:

$$\mathrm{TPBt}(x) = \sum\nolimits_{y\in\mathrm{Rx}} \sum\nolimits_{z\in\mathrm{Ry}} \mathrm{OD}(y,z,x) \frac{\mathrm{p}(z)}{\mathrm{Links}(y)} \quad (2)$$

wherein, TPBt(x) represents betweenness of the model; OD(y,z,x) means the shortest topological path from node y to node z through road section x within range of search radius R.

3.2.2. Linear Regression Model (OLS)

Hedonic Price Model (Hedonic)

The hedonic price model is a linear model function widely used in land value, price evaluation and prediction in real estate and other fields [40]. The specific function models include linear, semi-logarithmic, and dual logarithmic forms. However, the dual logarithmic model could more effectively reveal significant marginal utility of transaction price on house characteristics during purchase, so the simulation process was more practical and reasonable. Accordingly, the hedonic model used in this study consulted a dual logarithmic model [41] commonly used by previous scholars. The relationship between house price and each explanatory variable is listed as follows:

$$\ln Price(i) = \beta_0 + \sum_{i=1}^{i=n} \beta_n \ln x_{ni} + \varepsilon_i \quad (3)$$

Among them, *Price* in *Price*(i) represents house price, i means data in the i-th real estate, and *Price*(i) represents average price per square meter of the i real estate, including the sum of the price of the house itself and renovation costs; that is, the average transaction price of the i real estate. The n in x_{ni} is a characteristic factor, which represents the n-th characteristic influencing factor among many attributes; i indicates data in the i-th real estate, and x_{ni} is data performance of the n-th characteristic influencing factor in the i-th real estate. β_n is the non-standardized coefficient of the n-th relevant characteristic-influencing factor on house price; ε_i is a stochastic error term. In addition to the coefficients in the model, regression analysis also exposes strong or weak significance of major positive or negative explanatory variables that affect house prices.

Interactive Regression Model

In econometrics, when linear regression model is added with an interaction term, it becomes a regression equation model with special processing form. The interaction is the result of mutual effect of two or more influencing factors. To some extent, this method broadens interpretation angle and depth of explained variables being affected by different explanatory variables in the regression model. During research, additive interaction terms and multiplicative interaction terms were considered separately. However, after comparison of the significance of relevant data, a multiplicative interaction model with better fit and significance was selected as the research method for interaction. The specific calculation formula model is as follows:

$$y_i = \beta_i + \beta_{\mathrm{m}} x_{mi} + \beta_n x_{ni} + \beta_{mn} x_{mi} x_{ni} + \varepsilon_i \tag{4}$$

$$\frac{\partial y_i}{\partial_{xmi}} = \beta_m + \beta_{mn} x_{ni} \tag{5}$$

$$\frac{\partial^2 y_i}{\partial x_{mi} \partial x_{ni}} = \frac{\partial\left(\frac{\partial y_i}{\partial x_{mi}}\right)}{\partial x_{ni}} = \beta_{mn} \tag{6}$$

Find the partial derivative of x_{mi} for the first function to obtain second function formula above. In other words, after an interaction term is added to the original model, marginal effects of explanatory variables x_{mi} and y_i change; marginal effects originally depended on constants changing into that relying on explanatory variable x_{ni}; that is, the linear function equation model associated with x_{ni}. When $\beta_{mn} > 0$, the marginal effect of x_{mi} and y_i will increase as the variable x_{ni} increases. This change is called "synergy effect" in the marketing. On the contrary, when $\beta_{mn} < 0$, the marginal effect of x_{mi} will decrease with the rising of x_{ni}. Continuously find the partial derivative of x_{mi} for the second function to obtain third function formula above. β_{mn} is the "interaction effect" in the interaction model, which is specifically expressed as the relative effect of an explanatory variable on y_i and strength of the effect affected by another explanatory variable.

3.2.3. Spatial Regression Model

Spatial Error Model

The spatial error model specifically describes the influence of spatial nuisance dependence on explanatory variables in the space as a whole. Alternatively, it can be understood that a certain spatial disturbance will affect other spaces along with space benefits. The spatial error model is suitable to explain that spatial disturbance occurring in a certain space between regions will be transmitted to e adjacent space in the form of covariance, which is the result of random interaction in space. These effects are long-term and continuous, and will gradually decay with time or space factors. The calculation model is as follows:

$$\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon = \lambda W \varepsilon + \mu \\ u \sim \mathrm{N}(0, \sigma^2 \mathrm{I}) \end{cases} \tag{7}$$

wherein, λ represents spatial error correlation coefficient in spatial error model; W means spatial weight matrix with multiple spatial data points in spatial error model, which explains proximity relation of explanatory variables; and $\mathrm{w_{ij}}$ describes the correlation between individual error terms of the *j*-th and the *i*-th spatial data point.

Spatial Auto Regression Model

The space lag model means that a certain space is not only affected by the explanatory variables of the space, but also other similar spaces in the space. It aims to analyze the impact of spatial auto-correlation of house prices on house price itself. The spatial auto regression model is as follows:

$$\mathrm{Y} = \rho \mathrm{WY} + \mathrm{X}\beta + \varepsilon,\ \varepsilon \sim \mathrm{N}\left(0, \sigma^2 \mathrm{I}\right) \tag{8}$$

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \quad (9)$$

Among them, in the spatial lag model, Y is a matrix of explained variables; ρ represents the spatial effect coefficient; X means a matrix of explanatory variables; β represents vector of parameters. Similarly, W is a spatial matrix with multiple spatial data points; that is, to describe the mutual influence relationship between explanatory variables and the adjacent spatial explanatory variables. The n in the matrix means n spatial data points, and the matrix explains proximity relation of spatial regions between spatial data points. The spatial data points in each row represent a relationship set between a spatial unit and its proximity space, such as W_{11}~W_{1n}, indicating spatial proximity relation between spatial data point 1 and spatial data points 1, 2, . . . , n.

3.3. Data

This study mainly selected 1245 residential community samples from twelve county-level administrative regions under Fuzhou City, Fujian Province, China as the research object (Figure 2). In previous studies on house prices, Fuzhou was rarely used as a reference city due to its low international reputation. However, in recent years, Fuzhou has been in the development stage of urban rail construction, so it is quite meaningful to study the impact of rail station on house prices. Furthermore, the study results can guide and review the urban rail station under construction. Therefore, comprehensively considering the novelty and value of the research target, the study selected Fuzhou as the reference city.

Figure 2. Location and Perspective of Fuzhou City.

The data were all spatial market data from March 2021 in this study, and the dependent variable house price data are the average data in one month. Since other variables basically did not change within a month, the spatial cross-section data on 15 March 2021 are used as a representative. The selection of independent variables is the result of significant screening based on the experience of literature which may affect house prices. Average price was taken as basic unit of comparison, and large real estates with different residential properties were eliminated (such as villa estates), in order to avoid abnormal fluctuation of price of each house due to exquisite degree of decoration, north/south orientation, and elevator equipment, as well as other individualized and diversified factors [42].

Considering the realities of the study area, a brief introduction to the variables of the study is now carried out. The explained variable was represented by the average transaction price of residential communities, and information and data related to residential community were used as the explanatory variables that were divided into three categories: (1) self-characteristic information data and location information data, which covered the administrative area factor, urban location factor, nature of property rights, construction years, types of floors, plot ratio, greening rate, the level of property management fees, and educational resources (whether there were key primary and secondary schools) of residential communities. Since in Fuzhou, there are only the Second Ring Road and the

Third Ring Road in the urban area, and there is no First Ring Road, the entire urban space area is divided into three areas: the area within the Second Ring Road, the area from the Second Ring Road to the Third Ring Road, and the area outside the Third Ring Road. Among the variables, we used dummy variable 1 (CityLC1) and dummy variable 2 (CityLC2) of the urban space area to distinguish the three regions geographically. For the distinction of administrative space regions, we used the dummy variable (AdmRC) of whether it belongs to the city administrative region as the criterion. (2) Other external neighborhood explanatory variables included explanatory variables positively correlated with prices, such as shopping malls, grade 3 and first-class hospitals, scenic spots, green spaces, and important water bodies, and those negatively related to prices, namely funeral facilities, factories, gas stations and garbage dumps. Their influence on properties rested with the distance from residential communities. (3) The accessibility explanatory variables contained road accessibility variables (closeness, betweenness), rail accessibility variable (Euclidean distance to the closet subway station), and interaction variables between road accessibility variables and rail accessibility variables. Note: The price data comes from Anjuke official website database, and Ovital Map provides neighborhood and accessibility data. Please refer to figures below for specific explanatory variable details (Table 1, Figure 3) and road calculation results (Table 2, Figures 4 and 5).

Table 1. Descriptive statistics of the property dataset (N = 1245).

Variable	Description	Mean	Std. Dev.
Pri	House price (CNY/m^2)	26,526.8619	11,150.4492
	Explanatory variables		
Sub	Distance to the closest subway station of property (m)	3713.1051	8718.0624
NQ	Closeness of the closest road to property (C)	49.7878	10.3984
TP	Betweenness of the closest road to property (C)	80.0433	551.6053
Inter (Sub&NQ)	Interaction of SubS and NQ	-	-
Inter (Sub&TP)	Interaction of SubS and TP	-	-
	Control variables		
AdmRC	Dummy variable, 1 if the property in the urban district, 0 otherwise	0.8600	0.3440
CityLC1	Dummy variable, 1 if the property inside Second Ring Road, 0 otherwise	0.4000	0.4900
CityLC2	Dummy variable, 1 if the property inside Third Ring Road, 0 otherwise	0.7400	0.4410
TypeC	Dummy variable, 1 if the property belongs to commercial housing, 0 otherwise	0.8800	0.3310
AgeC	Dummy variable, 1 for a property built after 2000, 0 otherwise	0.8800	0.3210
HeightC	Dummy variable, 1 for a property with high-rise building and above, 0 otherwise	0.8200	0.3840
VolR	Volume Rate of community (C)	2.3835	1.1766
GreR	Greening rate of community (C)	0.3278	1.1932
ManF	Average management fee of community for each month (CNY/m^2)	1.1860	0.6561
PriS	Dummy variable, 1 for a property with high-quality primary school, 0 otherwise	0.0538	0.2257
MidS	Dummy variable, 1 for a property with high-quality middle school, 0 otherwise	0.0635	0.2439
Mal	Distance to the closest mall (m)	23,447.4507	6431.8477
Hos	Distance to the closest first-class Hospital at Grade 3 (m)	4794.6501	9618.9528
Sce	Distance to the closest scenic spot (m)	2167.6270	5769.8171
Lan	Distance to the closest landscape space (m)	2939.3625	8718.3601
Wat	Distance to the closest main water body (m)	1742.4330	4145.9548
Fur	Distance to the closest funeral facility (m)	4240.9892	5171.8461
Fac	Distance to the closest factory (m)	1892.3115	1878.8172
Pet	Distance to the closest petrol station (m)	1653.2890	3402.2372
Dum	Distance to the closest dump (m)	10,313.8575	4859.5494

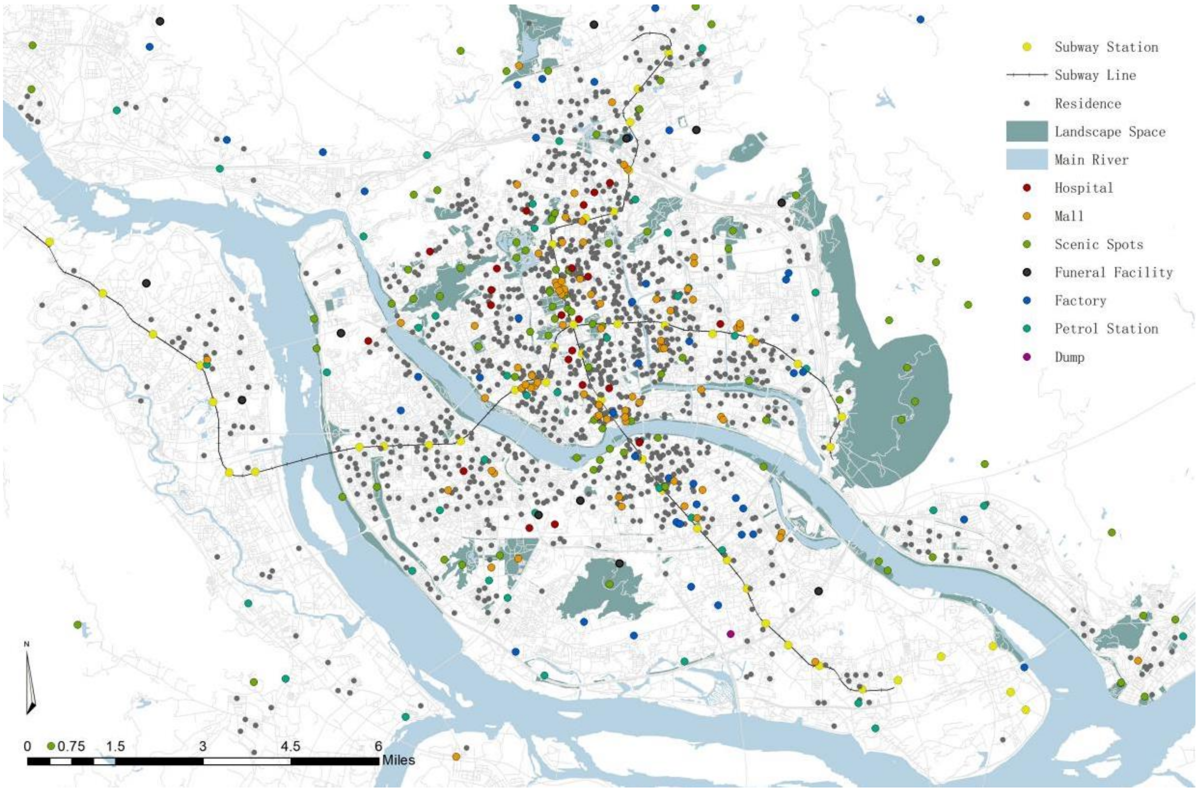


Figure 3. Schematic of the spatial distribution of variables.

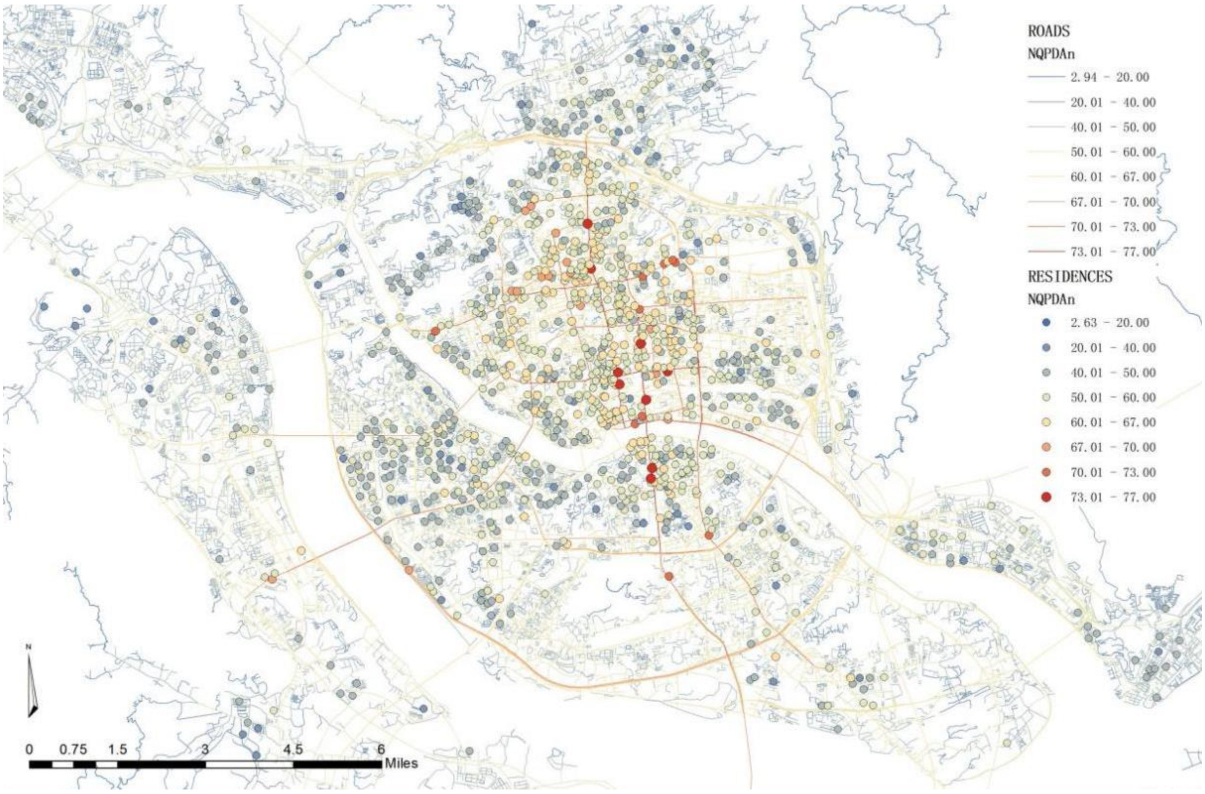


Figure 4. Result of connectivity.

Table 2. Statistics of the Space Syntax Analysis (N Roads = 56,228, N Residences = 1245).

Variable	**Range of Results**								**Objective**
NQ(NQPDAn)	0–20	20–40	40–50	50–60	60–67	67–70	70–73	73–77	
	579	8639	22,210	17,344	5844	699	546	367	Roads
	90	98	336	386	243	43	26	23	Residences
TP(TPBtAn)	0–50	50–150	150–350	350–600	600–1000	1000–1500	1500–2500	2500–4500	
	48,190	3470	1712	656	739	341	449	671	Roads
	998	93	41	18	27	9	16	43	Residences

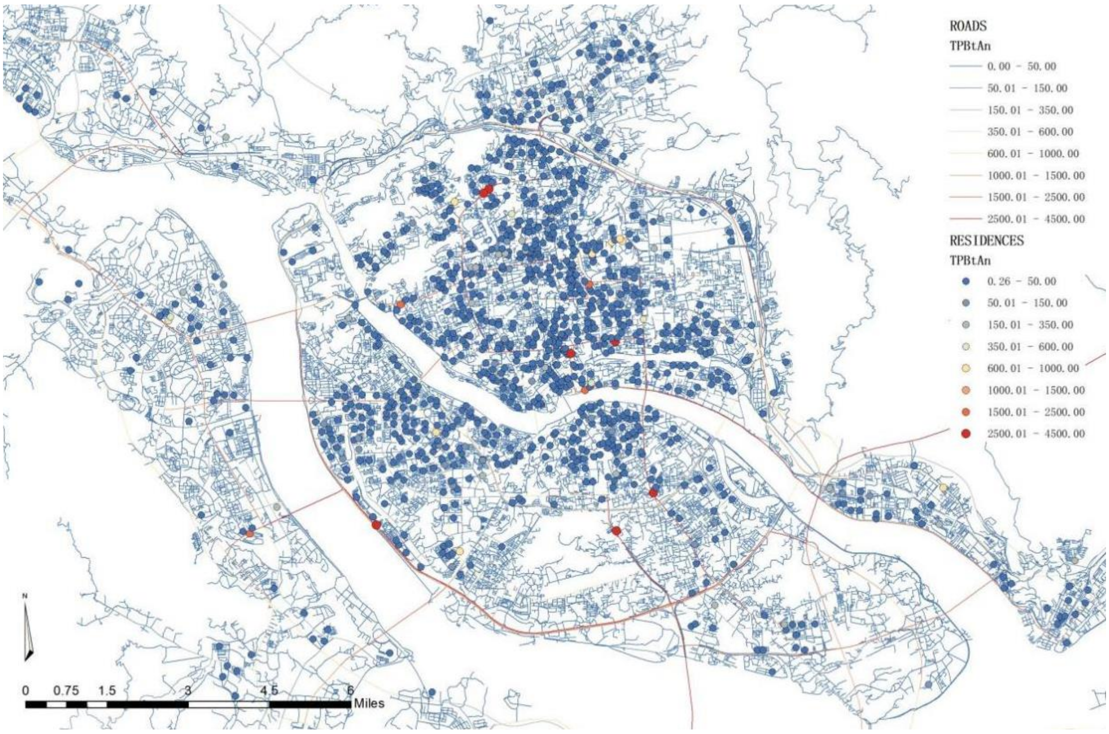


Figure 5. Result of carrying capacity.

4. Results

4.1. OLS Regression

After sorting out collected explanatory variables of the residential community, this study obtained a hedonic price model through regression, and then added interaction terms of road description variables and a rail distance variable for another linear regression of the model, acquiring a linear interactive regression model of house prices.

According to the regression equation of the interactive model, regression was generally good ($R^2 \geq 0.675$), and had the expected explanatory power.

From the perspective of significance, house price was significantly affected by explanatory variables (distance to subway stations, road network proximity), and the interaction of both. However, and the influence of road network betweenness and its interaction with distance to subway stations was insignificant.

The significant characteristics expressed by other control variables are as follows. Variables (locations in urban area, in second ring road and in the third ring road, average property fees, key primary schools and key middle schools nearby, distance to grade 3 and first-class hospital, distance to funeral facilities, distance to factories, distance to important water bodies and the distance to garbage disposal facility) exerted a significant impact on house prices. The impact of distance between residential communities and scenic spots on house price was generally correlated. However, the relationship between house prices with explanatory variables (nature of properties, building year, types of floors, greening rate, plot ratio, the distance to green space, distance to the closet business district, and the distance to the closet gas station) was insignificant.

This study conducted two robustness check analyses so as to verify the variables of distance to subway station, road network proximity, and the stability of the significance of their interaction. In the first check, some variables with low significance were removed and then regression fitting was performed again. In the second check, the natural logarithm of the road accessibility variable in the model was replaced by a constant, and a regression test was obtained similar to the previous regression, proving the reliability of above linear regression results. Please refer to Table 3 for overall results of the linear regression model and the test model.

Table 3. Regression results of the OLS.

Variable	Hedonic Price Model (Double-Log)			Econometrics Interaction Model		Robustness Checks Analysis 1		Robustness Checks Analysis 2	
	Coefficient	**T-Statistic**	**VIF**	**Coefficient**	**T-Statistic**	**Coefficient**	**T-Statistic**	**Coefficient**	**T-Statistic**
(Constant)	7.2100 **	22.1530		9.2980 **	−12.5480	9.3500 **	15.1270	7.9940 **	30.0280
Sub	−0.0370 **	−4.2930	2.7140	−0.2840 **	−3.6120	−0.2780 **	−4.1200	−0.1420 **	−5.2220
NQ	0.1350 **	3.5840	2.6550	−0.4570	−2.3700	−0.4410 **	−2.7830	−0.0140 **	−3.2220
TP	−0.0050	−1.2190	1.2560	−0.0020	−0.0560	–	–	0.0236 **	0.2900
Inter (Sub&NQ)	–	–		0.0640 **	3.1060	0.0620 **	3.5710	0.0020	4.0700
Inter (Sub&TP)	–	–		0.0000	0.1060	–	–	0.0000	−0.8190
AdmRC	0.3280 **	9.4740	3.4310	0.3220 **	9.3130	0.3320 **	10.7120	0.3350 **	9.7320
CityLC1	0.0940 **	5.1540	1.9260	0.1130 **	5.9040	0.1120 **	6.0670	0.1160 **	6.1160
CityLC2	0.2370 **	8.5300	3.6190	0.2460 **	8.8390	0.2480 **	9.1600	0.2440 **	8.8430
TypeC	0.0340	1.6750	1.0660	0.0320	1.5730	0.0320	1.6140	0.0320	1.5900
AgeC	0.0350	1.5320	1.2910	0.0340	1.4850	–	–	0.0370	1.6270
HeightC	0.0120	0.6740	1.1090	0.0150	0.8340	–	–	0.0150	0.8520
VolR	0.0170	0.9910	1.1160	0.0220	1.2910	–	–	0.0190	1.1170
GreR	0.0250	1.4460	1.2110	0.0230	1.3650	–	–	0.0250	1.4650
ManF	0.1730 **	12.0200	1.3550	0.1730 **	12.1020	0.1870 **	14.1220	0.1740 **	12.2140
PriS	0.2620 **	8.4200	1.1960	0.2690 **	8.6470	0.2580 **	8.4330	0.2740 **	8.8380
MidS	0.3470 **	11.1450	1.3620	0.3590 **	11.4890	0.3590 **	11.9610	0.3550 **	11.4180
Mal	0.0000	−0.0360	4.7840	−0.0010	−0.1290	–	–	0.0010	0.1180
Hos	−0.0420 **	−3.7510	2.5610	−0.0430 **	−3.8460	−0.0410 **	−3.7230	−0.0410 **	−3.6610
Sce	0.0170	1.6590	1.4360	0.0240	2.2940	0.0220	2.3150	0.0260 **	2.4990
Lan	−0.0070	−0.8460	1.6470	−0.0050	−0.6830	–	–	−0.001	−0.1650
Wat	−0.0560 **	−7.0010	1.7460	−0.0560 **	−7.0540	−0.060 **	−8.3450	−0.0570 **	−7.2110
Fur	0.0290	2.4000	1.4460	0.0450 **	3.4180	0.0430 **	3.4540	0.0450 **	3.4500
Fac	0.0420 **	4.2760	2.3960	0.0390 **	3.8630	0.0410 **	4.1170	0.0420 **	4.2210
Pet	0.0060	0.5660	3.9900	0.0070	0.6390	–	–	0.0110	1.1030
Dum	0.1530 **	9.5160	1.4580	0.0160 **	9.9000	0.1590 **	9.9750	0.1650 **	10.0710
Performance statistics									
R−squared	0.6750			0.6780		0.6760		0.6820	
Adjusted R-squared	0.6690			0.6710		0.6720		0.6750	

** Parameters are significant at the 1% level, * Parameters are significant at the 5% level.

4.2. Interactive Spatial Effect on Property Prices

Based on the linear interactive regression model established above, this study further took into account spatial errors in and spatial correlation factors of explained variables respectively into the model. In accordance with the spatial location distance of properties, a geographic spatial matrix was established to construct a spatial interactive regression model.

After consideration of the spatial correlation of errors and the spatial autocorrelation of explained variables, the spatial interactive regression model constructed had a slightly better regression fitting degree than the linear regression interaction model $\left(R^2_{SEM} = 0.685,\ R^2_{SAR} = 0.679\right)$. More importantly, significance and coefficient characteristics of interaction terms were consistent with those in the linear interactive regression model. This indicated impact of interaction between distance to subway stations and road network proximity on property in the linear model was also suitable for spatial relationship. The influence would not become weak with the changes in the relative spatial location of properties.

Similarly, for the purpose of testing stability of calculation data, it examined robustness of the spatial mode through two steps. The first was the change in the functional form of the fitting model. The previous dual-logarithmic model was replaced with a semi-logarithmic model that took the natural logarithm of the explained variable and the original value of the explanatory variable for regression fitting of the spatial model. Second was the reduction in the number of explanatory variables. Explanatory variables insignificant in the previous regression were removed and spatial linear regression analysis was conducted. Next, on the basis of the first test, dummy variables in all explanatory variables were abandoned again for another spatial regression analysis. After the test regression model was obtained, this study compared coefficient significance and characteristics of the tested spatial regression equation and the original spatial regression equation, and confirmed that the research results of the spatial interactive regression model were reliable. In Tables 4 and 5 results of the spatial regression model and robustness test are listed.

Table 4. Regression results and Robustness checks of the SLM and SEM.

Variable	Sem		Sem (Semi-Log) of Robustness Checks 1		Sar		Sar (Semi-Log) of Robustness Checks 1	
	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic
W_LnPri	-	-	-	-	0.1347	1.5407	−0.0248	−0.0834
Sub	−0.2928 **	−3.8009	0.0000	−1.8231	−0.2699 **	−3.4339	0.0000 *	−2.5322
NQ	−0.4983 **	−2.6462	0.0025 *	2.1127	−0.4294 *	−2.2396	0.0026 *	2.1581
TP	−0.0004	−0.0165	0.0000 *	−2.3156	−0.0016	−0.0584	0.0000 *	−2.2934
Inter (Sub&NQ)	0.066251 **	3.3110	0.0000	1.6816	0.0601 **	2.9446	0.0000 **	2.8612
Inter (Sub&TP)	0.0003	0.0884	0.0000	−0.3786	0.0004	0.1035	0.0000	−0.4333
CONSTANT	9.6992 **	13.2243	9.0414 **	19.3137	7.7664 **	6.1061	9.0699 **	3.0119
AdmRC	0.3001 **	8.9339	0.4346 **	9.1874	0.3250 **	9.5741	0.4774 **	10.0059
CityLC1	0.1073 **	5.7140	0.1927 **	10.1557	0.1145 **	6.0558	0.1958 **	10.1539
CityLC2	0.2448 **	9.0717	0.2095 **	6.7773	0.2445 **	8.9853	0.1985 **	6.3498
TypeC	0.0329	1.6800	0.0341	1.5753	0.0314	1.5882	0.0347	1.5785
AgeC	0.0357	1.6107	0.0107	0.4592	0.0342	1.5280	0.0066	0.2763
HeightC	0.0153	0.8881	0.0268	1.4315	0.0123	0.7086	0.0243	1.2715
VolR	0.0192	1.1588	0.0047	0.7610	0.0228	1.3608	0.0059	0.9446
GreR	0.0211	1.2754	−0.0053	−0.8925	0.0227	1.3608	−0.0042	−0.7065
ManF	0.1730 **	12.4211	0.1448	12.2464	0.1729 **	12.2917	0.1420	11.820 **
PriS	0.2650 **	8.7119	–	–	0.2715 **	8.8369	–	–
MidS	0.3620 **	11.9734	–	–	0.3544 **	11.5991	–	–
Mal	−0.0079	−0.9287	0.0000 **	−2.8338	0.0006	0.0666	0.0000	−1.4774
Hos	−0.0422 **	−3.8661	0.0000 **	−6.1241	−0.0441 **	−4.0080	0.0000 **	−7.2118
Sce	0.0162	1.5646	0.0000 **	3.5900	0.0271 **	2.5870	0.0000 **	4.3866
Lan	−0.0044	−0.6334	0.0000*	2.0285	−0.0066	−0.9347	0.0000	1.8612
Wat	−0.0612 **	−7.4480	0.0000 **	−6.3291	−0.0523 **	−6.1924	0.0000 **	−3.7617
Fur	0.0372 **	2.8948	0.0000 **	3.1900	0.0460 **	3.5522	0.0000 **	3.4413
Fac	0.0345 **	3.5203	0.0000 **	4.0072	0.0389 **	3.9474	0.0000 **	3.6615
Pet	−0.0046	−0.4271	0.0000 **	−2.9767	0.0112	1.0472	0.0000 **	2.9613
Dum	0.1671 **	10.5393	0.0000 **	12.6471	0.1568 **	9.6418	0.0000 **	11.4135
LAMBDA	0.8652 **	10.5974	0.9848 **	93.0953	–	–	–	–
Performance statistics								
R−squared	0.6850		0.6147		0.6791		0.6026	
AIC	−169.2720		82.9515		−148.2160		114.1090	
BIC	−35.9726		205.9970		−9.7901		242.2810	

** Parameters are significant at the 1% level, * Parameters are significant at the 5% level.

Table 5. Robustness checks results of the SLM and SEM.

Variable	SEM of Robustness Checks 2		SEM of Robustness Checks 3		SAR of Robustness Checks 2		SAR of Robustness Checks 3	
	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic	Coefficient	T-Statistic
W_LnPri	-	-	-	-	0.0443	0.5562	0.0883	0.9732
Sub	−0.2691 **	−4.1319	0.0604	0.8377	−0.2401 **	−3.5934	0.1006	1.3646
NQ	−0.4437 **	−2.8817	0.3702 *	2.1839	−0.3837 *	−2.4396	0.4503 **	2.6119
TP	—	—	—	—	—	—	—	—
Inter (Sub&NQ)	0.0598 **	3.5908	−0.0291	−1.5875	0.0527 **	3.0846	−0.0391 *	−2.0842
Inter (Sub&TP)	-	-	-	-	-	-	-	-
CONSTANT	9.6050 **	15.6409	7.4893 **	10.8421	8.8098 **	7.9145	6.1941 **	4.9134
AdmRC	0.3047 **	10.2107	0.3749 **	11.6312	0.3197 **	10.5517	0.3917 **	11.9938
CityLC1	0.1041 **	5.8734	-	-	0.1057 **	5.8950	-	-
CityLC2	0.2522 **	9.4950	-	-	0.2533 **	9.4373	-	-
TypeC	-	-	-	-	-	-	-	-
AgeC	-	-	-	-	-	-	-	-
HeightC	-	-	-	-	-	-	-	-
VolR	-	-	-	-	-	-	-	-
GreR	-	-	-	-	-	-	-	-
ManF	0.1895 **	14.6114	0.1601 **	10.8034	0.1899 **	14.4630	0.1613 **	10.7307
PriS	0.2538 **	8.4647	-	-	0.2544 **	8.3863	-	-
MidS	0.3529 **	12.0357	-	-	0.3490 **	11.7549	-	-
Mal	-	-	-	-	-	-	-	-
Hos	−0.0397 **	−3.6868	−0.1437 **	−14.8864	−0.0380 **	−3.5020	−0.1422 **	−14.6306
Sce	-	-	-	-	-	-	-	-
Lan	-	-	-	-	-	-	-	-
Wat	−0.0617 **	−8.3078	−0.0514 **	−6.0215	−0.0560 **	−7.3713	−0.0454 **	−5.2331
Fur	0.0344 **	2.7474	0.0754 **	5.3473	0.0444 **	3.5099	0.0859 **	6.0436
Fac	0.0360 **	3.7026	0.0517 **	4.8245	0.0418 **	4.2664	0.0566 **	5.2641
Pet	-	-	-	-	-	-	-	-
Dum	0.1653 **	10.5287	0.0517 **	4.8245	0.1560 **	9.6766	0.1117 **	6.3729
LAMBDA	0.8622 **	10.3570	0.8312 **	8.3333	-	-	-	-
Performance statistics								
R-squared	0.6815		0.5729		0.6744		0.5663	
AIC	−177.6170		179.2460		−151.8840		197.0250	
BIC	−100.7140		235.6420		−69.8536		258.5470	

** Parameters are significant at the 1% level, * Parameters are significant at the 5% level.

After the spatial regression interaction equation was acquired, functional images were used to represent the impact of interaction on house price in different regions in order to reveal the impact of interaction on explained variables in different spatial regions (heterogeneous regions). According to the administrative districts of residential communities, data of existing properties were divided into five sub-sets (Gulou District, Taijiang District, Cangshan District, Jin'an District, Minhou County) along the existing subway lines. The spatial classification was shown in Figure 6. In functional images, sample residential communities in Fuzhou were based on the average level of all 1245 samples, and variable values of sample residential communities were subject to an average level in the administrative districts (continuous variables take the average value in the area, and dummy variables are determined as 0 or 1 by rounding the average value; see Table 6). As the value method of control variables, specific function images of each area sample are shown in Figure 7.

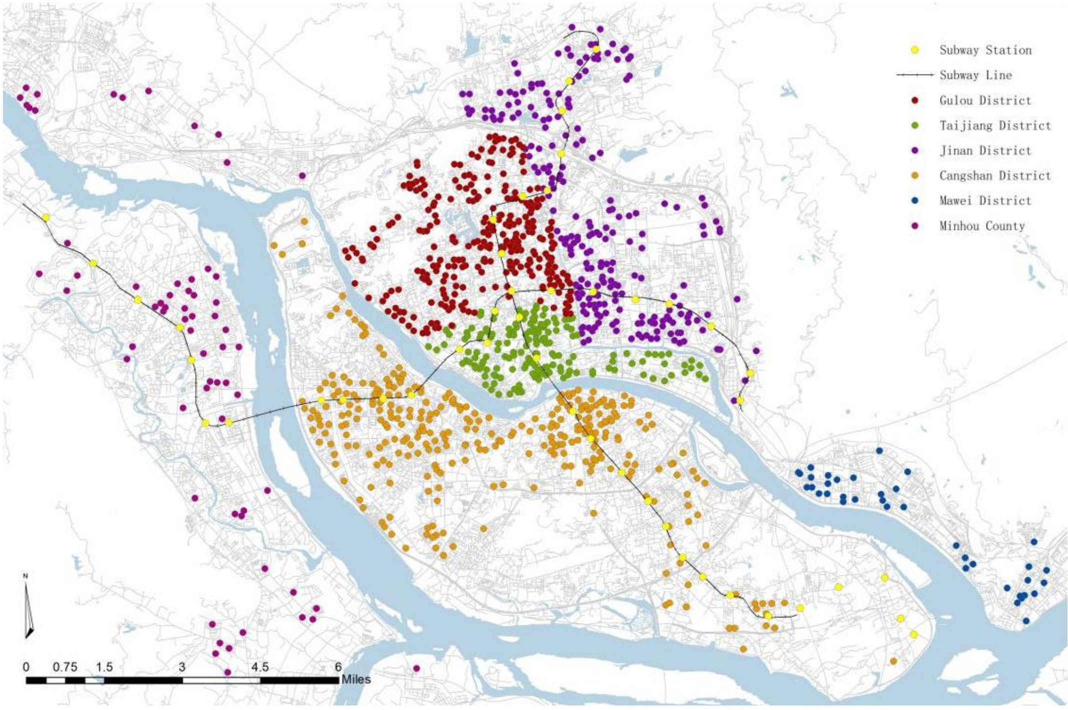


Figure 6. Administrative Division of Residential Space.

Table 6. Mean Statistics of Sample Properties (N = 1245).

Variable/District	Fuzhou	Gulou	Taijiang	Cangshan	Jin'an	Minhou
Explained variable						
Pri	-	-	-	-	-	-
Explanatory variables						
Sub	0–5000	0–5000	0–5000	0–5000	0–5000	0–5000
NQ	0–80	30–75	25–80	25–75	30–75	30–70
Control variables						
AdmRC	1	1	1	1	1	0
CityLC1	0	1	1	0	0	0
CityLC2	1	1	1	1	1	0
TypeC	1	1	1	1	1	1
AgeC	1	1	1	1	1	1
HeightC	1	1	1	1	1	1
VolR	2.3835	2.3064	2.4354	2.3218	2.5017	2.3165
GreR	0.3278	0.2483	0.2584	0.3860	0.3975	0.2584
ManF	1.1860	1.0233	1.0661	1.1564	1.1997	1.6264
PriS	0	0	0	0	0	0
MidS	0	0	0	0	0	0
Mal	23,447.4507	702.7686	615.6014	1100.3267	958.2673	3680.7538
Hos	4794.6501	857.1616	1101.4230	2081.3574	1896.6576	7838.2267
Sce	2167.6270	674.0358	638.1203	1392.7767	924.3638	2593.8097
Lan	2939.3625	257.0721	217.0599	552.1107	445.0455	4836.5047
Wat	1742.4330	1143.0556	873.4067	1059.9912	1366.6342	1468.2550
Fur	4240.9892	3725.6462	3430.5961	2465.5883	3029.2090	3730.6051
Fac	1892.3115	1424.2365	1119.5498	1622.8429	966.6642	4653.3776
Pet	1653.2890	1181.7450	1076.5128	1142.1527	1057.8335	1866.0762
Dum	10,313.8575	11,385.0439	7810.7964	6890.3944	11,028.7819	13,013.5057
TP	80.0433	96.7403	105.5944	111.0925	38.5952	77.1311

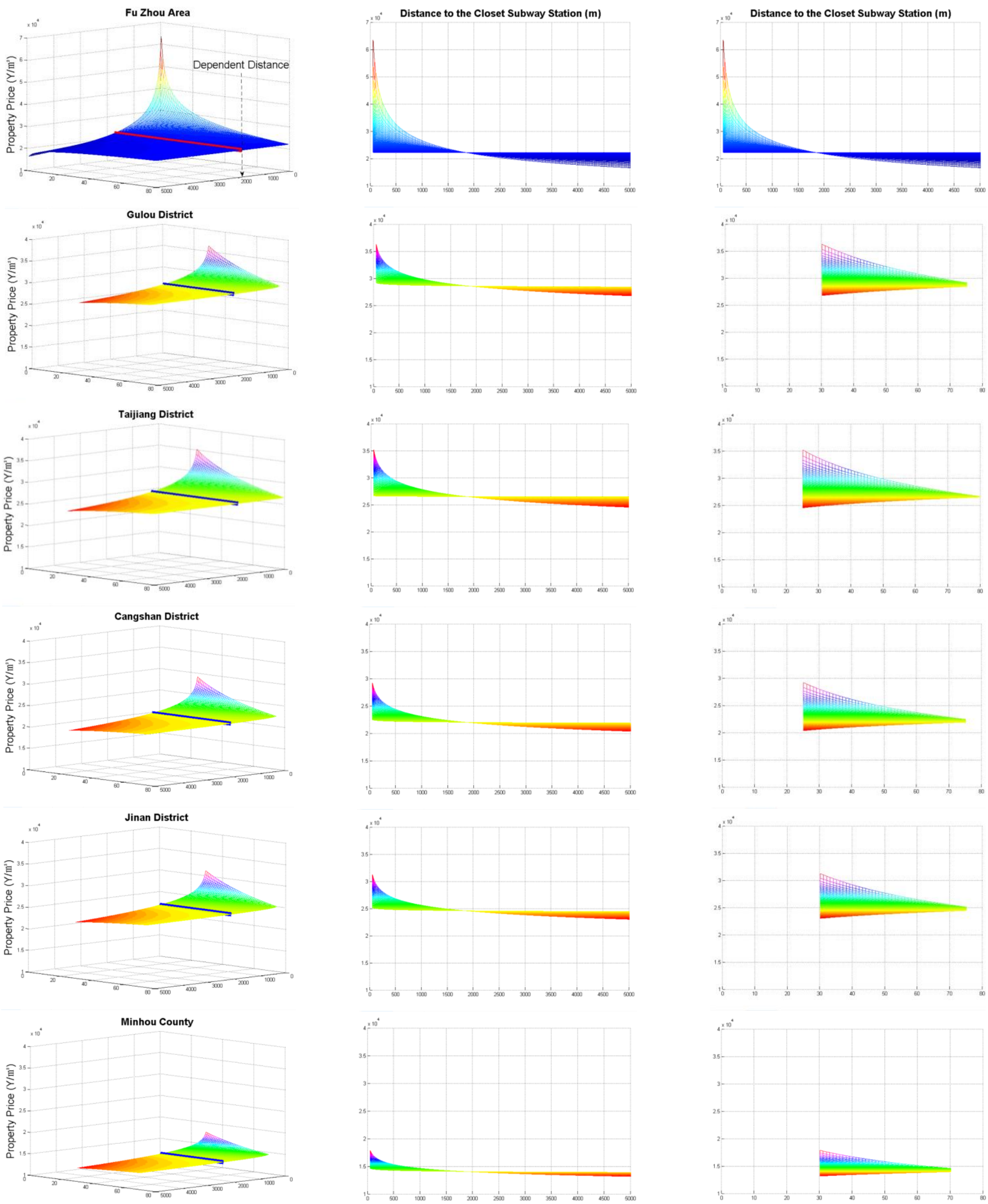


Figure 7. The impact of accessible variables on property price.

5. Discussions and Conclusions

5.1. Discussions of Findings

Findings of the analysis are drawn as per the above data table and coefficient interpretation analysis of images:

(1) This study classified the explanatory variables of properties and evaluated their significance, positive and negative effects. Among them, there was a significant negative correlation between the distance to subway stations and house price, or in other words, subway transportation produced a significant impact on house price. Traditional research showed that house price was strongly positively affected by house advantage factors such as the regional administrative nature, urban centrality, greening rate, property service level, key elementary and middle school districts, convenience of hospitals, and convenience to access water in the landscape. Apart from that, NIMBY factors (distance to funeral service sites, factories, and garbage dumps) produced a negative impact, and the influencing factors were more significant. A significant margin was discovered between property nature and construction age with house price, but insignificantly correlated with building height, plot ratio, and greening rate. The external neighborhood variable presented in the study is the distance. Taking the hospital variable as an example, the greater the distance between the residence and the hospital facilities, the lower the medical convenience would be, which also will make house price lower, so the coefficient of this variable is negative. On the contrary, the farther the factory facilities and funeral facilities are from the house, the better people's inner experience and environmental feelings, which will rise the house price, so the correlation coefficient is positive. The above findings are similar to the related research of previous scholars. For some dummy variables of the self-characteristic, such as administrative nature and geographic regional centrality, in addition to being able to see the impact direction caused by the positive or negative coefficients, the coefficients also reflect the strength of the impact. The influence intensity of the city administrative location variable (AdmRC) is stronger than that of the city geographic location variable (CityLC1 and CityLC2). This reflects that buyers' recognition of administrative factors is stronger than that of geographic central locations. In the comparison of the location of the urban geographic center, the influence intensity of the location in the third ring is greater than that of the location in the second ring. This reflects that the third ring road is stronger than the second ring road in terms of the degree of homebuyers' recognition of the convenience attached to the city's geographic centrality. This means that moving your home from outside the Third Ring Road to inside will cost more compared to the Second Ring Road.

(2) The better the road network proximity where properties were located, the higher the price, but the influence of road network betweenness was far from significant. Without the interactive impact of road accessibility and rail accessibility, it is indicated from Table 3 that road network proximity was significantly positively correlated with house prices, which proved the fact that buyers spent more on houses with better road connectivity. An insignificant negative correlation was found between road network betweenness and house prices; that is, priced declined due to the carrying capacity of roads. This was not completely consistent with conclusions that road network betweenness was negatively correlated with house prices in previous literature. After interaction between road network betweenness and the distance to subway stations, the influence of interactive variables was still insignificant.

(3) House price was much higher under the impact of rail accessibility than road connectivity that, however, buffered the extent of the influence of subway traffic on house prices. When interaction between distance to subway stations and road network proximity was considered in the model, it was understood from standard coefficients of the fitting equation that demand for subway traffic was higher than road connectivity. According to the analysis of partial derivative ($\frac{\partial \mathrm{LnPri}}{\partial \mathrm{LnSub}} = \beta_{\mathrm{Sub}} + \beta_{\mathrm{Inter(Sub\&NQ)}} \times \mathrm{LnNQ}$) of distance from subway station to the house, and the observation in Figure 7, road network proximity dramatically regulated the extent of influence of distance to subway station on house price. Generally, the better the road connectivity of urban streets where properties are located, the higher the road accessibility compensation for properties, and the lower the house price rise under the influence of subway stations.

Moreover, the derivative formula showed that road network proximity failed to make the partial derivative of distance to subway station to house price zero within the extreme value range of Fuzhou. Therefore, the influence of subway stations on house price was permanent.

(4) From the perspective of the overall average of Fuzhou, when properties were 1800 m away from subway station, the price would not change due to changes in road network proximity that could not only measure advantages of roads, but also represent unfavorable influencing factors behind roads. In Fuzhou, urban land transportation is composed of road transportation and subway transportation, and water and air transportation within the city can be basically ignored. Therefore, in the interactive model, road network proximity was used to find the partial derivative of house price; when the partial derivative was equal to zero ($\frac{\partial \text{LnPri}}{\partial \text{LnNQ}} = \beta_{\text{NQ}} + \beta_{\text{Inter(Sub\&NQ)}} \times \text{LnSub}$), it calculated that a distance of 1800 m to the subway station (shown by the cut-off line in Figure 7) was the critical range that people relied on subway stations. Beyond this range, people needed to pay additional expenses in house purchase. Further study indicated that road connectivity played bidirectional roles. Specifically, when properties were close to subway stations so that people could meet most of the daily travel, house prices would decline partially in accessibility variables due to noise and congestion in areas with better road network connectivity.

(5) The absolute value of the influence on price in different areas was different in interaction between rail accessibility and road connectivity of road traffic accessibility. According to Figure 7, due to different control variables in different districts, the house price was far from the same caused by interaction. In Fuzhou, the maximum price difference was CNY 40,000/km^2. In administrative districts, the largest difference was in Taijiang District, with a price difference of over CNY 10,000/km^2, while Minhou County had the smallest difference of more than CNY 5000/km^2.

(6) Under the background of urbanization, with other conditions remaining unchanged, during the construction and improvement of the road network, when the distance of the spatial direct line to a subway station on the same road is larger, the difference in house prices will be larger. Moreover, after road construction, the greater the proximity difference, the more significant the difference in house prices. This finding was further proven in Figure 8.

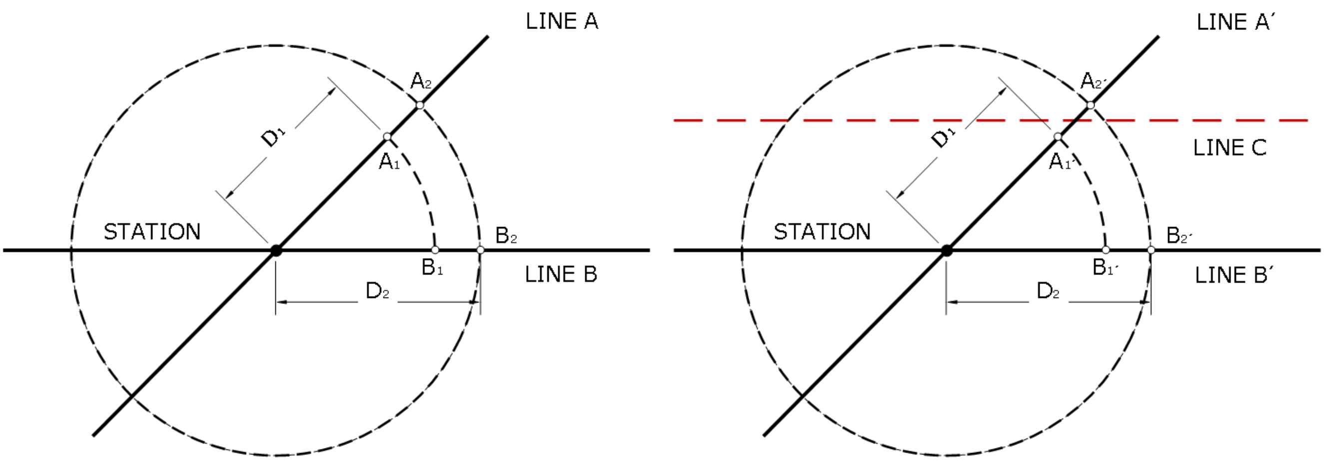


Figure 8. The impact of accessible variables on house price.

wherein

A_1 and A_2 represent different houses on line A

B_1 and B_2 represent different houses on line B

D_1 and D_2 ($D_1 \neq D_2 \geq 0$) represent the distance from properties to the closet rail station

P_n represents house price of n and ΔP_N represents the difference in house prices on line N

If $D_1 < D_2$ and other variables of A_1 and A_2 are the same, then $\Delta P_A\left(P_{A_1} - P_{A_2}\right) > 0$.

If closeness (line A) > closeness (line B), $D_1 < D_2$ and other variables are the same, then we can obtain

$$\left|\Delta P_A\left(P_{A_1} - P_{A_2}\right)\right| < \left|\Delta P_B\left(P_{B_1} - P_{B_2}\right)\right|$$

If $D_1 < D_2$ and add line C to intersect line A to make line A(B) turn into line $A'\left(B'\right)$ and $\Delta\text{Closeness}\left(\text{line } A' - \text{line } A\right) > \Delta\text{Closeness}\left(\text{line } B' - \text{line } B\right) > 0$ with others are the same, then

$$\left|\Delta P_A\left(P_{A_1} - P_{A_2}\right)\right| > \left|\Delta P_A{}'\left(P_{A_1}{}' - P_{A_2}{}'\right)\right|$$
$$\left|\Delta P_B\left(P_{B_1} - P_{B_2}\right)\right| > \left|\Delta P_B{}'\left(P_{B_1}{}' - P_{B_2}{}'\right)\right|$$
$$\left|\Delta P_A\left(P_{A_1} - P_{A_2}\right)\right| - \left|\Delta P_A{}'\left(P_{A_1}{}' - P_{A_2}{}'\right)\right| > \left|\Delta P_B\left(P_{B_1} - P_{B_2}\right)\right| - \left|\Delta P_B{}'\left(P_{B_1}{}' - P_{B_2}{}'\right)\right|$$

5.2. Discussions of Limitation

This study verifies how house price is affected by rail accessibility and road accessibility, as well as the interaction of both through the linear regression interaction model and spatial interaction term model, and reveals the mutual regulating role of accessibility variables on house prices. However, there are some shortcomings. Specifically, it is a complicated process to analyze the impact of interaction on house prices. Furthermore, cities are abundant in specific spatial location functions, and unique, which will change the extent of the influence of the interaction on house prices [43]. Unfortunately, limited by data source, personal abilities, and research time, this study fails to analyze problems comprehensively as follows.

First of all, the sample data selected are too narrow to fully reflect realities in Fuzhou. The findings should be verified further, and this decreases the fitting degree of the equation, affecting the coefficient of the regression equation thereby. In future research, continuous efforts should be made to collect more sample data and improve comprehensiveness of explanatory variables of house prices, so as to come to a more convincing finding with stronger explanatory power.

Then, when the hedonic price model is introduced to analyze house prices, characteristic explanatory variables omit relevant information and have deficiencies, which will reduce the accuracy of hedonic price model in decomposing real prices [44]. As a result, future scholars should consider more components that affect house prices, and split original components elaborately, in order to present a more realistic hedonic price model.

Next, the spatial interactive regression model is promoted because it is related to the spatial autocorrelation of explained variables and spatial disturbance correlation in error terms, without considerations on the spatial autocorrelation of all explanatory variables. In other words, the spatial Durbin model (SDM) is not mentioned or promoted over the general nested space model (GNS) to analyze the applicability of interaction between explanatory variables on the impact of house prices.

Finally, research findings are drawn based on a general traffic development law of cities — construction of subway traffic is slower than ideal construction speed. However, in reality, construction and development speed is diverse for different construction parts of cities, so, this implies that research findings should be analyzed dialectically. At the same time, it is hoped more scholars can study related fields to propose more suggestions and thoughts for model innovation and improvement, and supplement interaction theories of existing house price regression models.

5.3. Conclusions of Implication

In accordance with research purpose and results, this study will explain how to apply these research findings for government, individuals and scholars.

5.3.1. Policy Implications

In terms of third finding, it is recommended that governments should consider balanced urban development, the benefits of subway transportation, and fair transportation when planning subway transportation routes and subway stations. Apart from playing the role of organization, subway lines can stimulate the development of surrounding areas, including economic benefits, infrastructure, and guidance of population flow. If subway lines are arranged along the road network with high proximity, imbalance of this orientation will deteriorate, so that rising of spatial value of subway traffic declines. In summary, Fuzhou should allocate subway transportation lines and subway stations in urban areas with relatively low road network proximity. This will promote the balanced development of cities and provide urban residents with living spaces with better accessibility under low purchase costs. In addition the government can formulate targeted tax policies for areas near the built rail stations. By taking advantage of the population and economic vitality brought by rail stations, the investment in the previous transportation construction is recovered, which can form a virtuous economic cycle and promote urban construction and development.

5.3.2. Personal Implications

According to the fifth finding, when $D_1 > D_2 \geq 0$, then $\Delta P_A\left(P_{A_1} - P_{A_2}\right) >$, and $\Delta P'_A\left(P'_{A_1} - P'_{A_2}\right) > 0$; that is, the above formula (a) can be regarded as a transformation of $P_{A_1} - P_{A_2} > P'_{A_1} - P'_{A_2}$, so ${P_{A_2}}' - P_{A_2} > {P_{A_1}}' - P_{A_1}$ is obtained. Similarly, ${P_{B_2}}' - P_{B_2} > {P_{B_1}}' - P_{B_1}$ is acquired according to $P_{B_1} - P_{B_2} > {P_{B_1}}' - P_{B_2}$. Therefore, the following suggestions are given to buyers: in most cases, when road construction is faster than subway network construction; that is, similar to the above simulation of adding line C, only a single variable (distance between the closet subway station and house) can be considered. Buyers should purchase properties far away from subway stations in light of existing budgets, which is hopeful to improve the potential appreciation of properties. On the other hand, according to the overall level of Fuzhou, people relying more on subway transportation should purchase residential communities within 1800 m of subway stations. Otherwise, people have to pay more for housing expenses because they must rely on roads. Based on the above judgments, home buyers may need to combine their own economic budget and traffic preferences to determine the distance between the purchased residence and the rail station.

In the other hand, under the circumstances of same distance to subway stations but different road network proximity, in what kind of residential communities on road should one buy a house, in order to ensure the highest market value of properties held by buyers in the future? This issue should be discussed under different cases. In Fuzhou, when the subway station is within 1800 m of the residence, the price of properties on roads with smaller increases in road network proximity will be higher than that with significant improvement in road network proximity. The opposite is true beyond 1800 m; in other words, on roads with a smaller increase in road network proximity, the price will be lower than that with significant improvement in road network proximity. In summary, under the same distance to a subway station, residential communities closer to subway stations (within dependent distance) should be built on roads without much road development space, while those in the distance (beyond the dependent distance) are recommended to be constructed on streets with sufficient space for reserve roads construction. According to the above suggestions, home buyers will be able to obtain more residential investment income in the future.

5.3.3. Scholar Implications

The findings of this study partly explain the differences in the findings of previous scholars on similar issues, and also provide ideas for the academic community to study related issues. Previously, many scholars, such as Xin Wei et al. [45] and Zhengyi Zhou et al. [46], respectively studied the regional differences in the impact rail stations on house prices in Chengdu and Shanghai, and gave reasons for the differences. However, the inherent differences of the location variables discussed in their study are more likely to be caused by the different road network accessibility attached to the location variables. This study not only corroborates the authenticity of previous scholars' research, but also reveals the more essential reasons behind the location variables and improves the theory of the difference in the impact of subway stations on house prices. Moreover, researchers can use similar research ideas to study the differences in the impact of subway stations on house prices in the future, by building interactive regression models for other linear variables.

Since the development of nonlinear models, the use of this method to discuss various problems has been very extensive, both in the field of natural sciences and social sciences. The model can better explain the correlation between multiple variables, optimizing the incompleteness of variable classification discussions, simplifying the complexity, and promoting disciplinary exchanges. Therefore, the discussion of nonlinear models can also be extended to more discussions on price research [47]. In addition, transportation factors are affecting our lives in a more diverse and subtle way, so other scholars are called on to explore the development of transportation systems from the perspective of sustainable optimization in the future [48].

Author Contributions: Conceptualization, K.C. and H.L.; methodology, K.C., H.L. and L.L.; software, K.C. and L.L.; validation, K.C., Y.L. and Z.L.; formal analysis, K.C. and Y.L.; investigation, L.L. and Y.L.; resources, L.L.; data curation, L.L., L.T. and A.W.; writing—original draft preparation, K.C. and Y.L.; writing—review and editing, K.C.; visualization, K.C.; supervision, Y.-J.C.; project administration, L.L.; funding acquisition, L.L. and T.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [Distinguished Youth Program of Fujian Agriculture and Forestry University] grant number [XJQ201932] and [Interdisciplinary Integration Guidance Project of College of Landscape Architecture, Fujian Agriculture and Forestry University] grant number [YSYL-xkjc-3].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, L.; Chau, K.W.; Wang, X. Are low-end housing purchasers more willing to pay for access to public services? Evidence from China. *Res. Transp. Econ.* **2019**, *76*, 100734. [CrossRef]
2. Yang, L.; Wang, B.; Zhang, Y.; Ye, Z.; Wang, Y.; Li, P. Willing to pay more for high-quality schools? A hedonic pricing and propensity score matching approach. *Int. Rev. Spat. Plan. Sustain. Dev.* **2018**, *6*, 45–62.
3. Jia, J.; Zhang, X.; Huang, C.; Luan, H. Multiscale analysis of human social sensing of urban appearance and its effects on house price appreciation in Wuhan, China. *Sustain. Cities Soc.* **2022**, *81*, 103844. [CrossRef]
4. Yang, L.; Wang, B.; Zhou, J.; Wang, X. Walking accessibility and property prices. *Transp. Res. Part D Transp. Environ.* **2018**, *62*, 551–562. [CrossRef]
5. Tan, R.; He, Q.; Zhou, K.; Xie, P. The Effect of New Metro Stations on Local Land Use and Housing Prices: The Case of Wuhan, China. *J. Transp. Geogr.* **2019**, *79*, 102488. [CrossRef]
6. Cengiz, E.C.; Çelik, H.M. Investigation of the Impact of Railways on Housing Values; the Case of Istanbul, Turkey. *Int. J. Transp. Dev. Integr.* **2019**, *3*, 295–305. [CrossRef]
7. Law, S. Defining Street-based Local Area and Measuring its Effect on House Price Using a Hedonic Price Approach: The case Study of Metropolitan London. *Cities* **2017**, *60*, 166–179. [CrossRef]
8. Liang, J.; Koo, K.M.; Lee, C.L. Transportation infrastructure improvement and real estate value: Impact of level crossing removal project on housing prices. *Transportation* **2021**, *48*, 2969–3011. [CrossRef]
9. Lin, J.J.; Hwang, C.H. Analysis of Property Prices before and After the Opening of the Taipei Subway System. *Ann. Reg. Sci.* **2004**, *38*, 687–704. [CrossRef]
10. Yang, L.; Chau, K.W.; Chu, X. Accessibility-based premiums and proximity-induced discounts stemming from bus rapid transit in China: Empirical evidence and policy implications. *Sustain. Cities Soc.* **2019**, *48*, 101561. [CrossRef]
11. Medda, F. Land value capture finance for transport accessibility: A review. *J. Transp. Geogr.* **2012**, *25*, 154–161. [CrossRef]

12. Medda, F. Evaluation of Value Capture Mechanisms as a Funding Source for Urban Transport: The Case of London's Crossrail. *Procedia-Soc. Behav. Sci.* **2012**, *48*, 2393–2404.
13. Lee, C.L.; Locke, M. The effectiveness of passive land value capture mechanisms in funding infrastructure. *J. Prop. Investig. Financ.* **2021**, *39*, 283–293. [CrossRef]
14. Yang, L.; Liang, Y.; Zhu, Q.; Chu, X. Machine learning for inference: Using gradient boosting decision tree to assess non-linear effects of bus rapid transit on house prices. *Ann. GIS* **2021**, *27*, 273–284. [CrossRef]
15. Ridker, R.G.; Henning, J.A. The Determinants of Residential Property Values with Special Reference to Air Pollution. *Rev. Econ. Stat.* **1967**, *49*, 246–257. [CrossRef]
16. Shyr, O.; Andersson, D.E.; Wang, J.; Huang, T.; Liu, O. Where Do Home Buyers Pay Most for Relative Transit Accessibility? Hong Kong, Taipei and Kaohsiung Compared. *Urban Stud.* **2013**, *50*, 2553–2568. [CrossRef]
17. Cui, N.; Gu, H.; Shen, T.; Feng, C. The Impact of Micro-Level Influencing Factors on Home Value: A Housing Price-Rent Comparison. *Sustainability* **2018**, *10*, 4343. [CrossRef]
18. Zou, C.; Tai, J.; Chen, L.; Che, Y. An Environmental Justice Assessment of the Waste Treatment Facilities in Shanghai: Incorporating Counterfactual Decomposition into the Hedonic Price Model. *Sustainability* **2020**, *12*, 3325. [CrossRef]
19. Hitaj, C.; Lynch, L.; McConnell, K.E.; Tra, C.I. The Value of Ozone Air Quality Improvements to Renters: Evidence from Apartment Building Transactions in Los Angeles County. *Ecol. Econ.* **2018**, *146*, 721. [CrossRef]
20. Wang, J.; Lee, C.L. The value of air quality in housing markets: A comparative study of housing sale and rental markets in China. *Energy Policy.* **2022**, *160*, 112601. [CrossRef]
21. Agarwal, S.; Rengarajan, S.; Sing, T.F.; Yang, Y. School allocation rules and housing prices: A quasi-experiment with school relocation events in Singapore. *Reg. Sci. Urban Econ.* **2016**, *58*, 56. [CrossRef]
22. Agarwal, S.; Koo, K.M. Impact of electronic road pricing (ERP) changes on transport modal choice. *Reg. Sci. Urban Econ.* **2016**, *60*, 11. [CrossRef]
23. Xu, T. Heterogeneity in Housing Attribute Prices: A study of the Interaction Behaviour between Property Specifics, Location Coordinates and Buyers' Characteristics. *Int. J. Hous. Mark. Anal.* **2008**, *1*, 166–181. [CrossRef]
24. Currie, J.; Davis, L.; Greenstone, M.; Walker, R. Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings. *Am. Econ. Rev.* **2015**, *105*, 678–709. [CrossRef] [PubMed]
25. Im, J.; Hong, S.H. Impact of a New Subway Line on Housing Values in Daegu, Korea: Distance from Existing Lines. *Urban Stud.* **2018**, *55*, 3318–3335. [CrossRef]
26. Mayer, T.; Trevien, C. The impact of urban public transportation evidence from the Paris region. *J. Urban Econ.* **2017**, *102*, 1–21. [CrossRef]
27. Zheng, J.; Liu, H. The Impact of URRT on House Prices in Shenzhen. *J. China Railw. Soc.* **2005**, *27*, 11–18. (In Chinese)
28. Zhang, X.; Jiang, Y. An Empirical Study of the Impact of Metro Station Proximity on Property Value in the Case of Nanjing, China. *Asian Dev. Policy Rev.* **2014**, *2*, 61–71. [CrossRef]
29. Almosaind, M.A.; Dueker, K.J.; Strathman, J.G. Light-rail Transit Stations and Property Values: A Hedonic Price Approach. *Transp. Res. Rec.* **1993**, *1400*, 90–94.
30. Wen, H.; Gui, Z.; Tian, C.; Xiao, Y.; Fang, L. Subway Opening, Traffic Accessibility, and Housing Prices: A Quantile Hedonic Analysis in Hangzhou, China. *Sustainability* **2018**, *10*, 2254. [CrossRef]
31. Armstrong, R.J.; Rodriguez, D.A. An evaluation of the accessibility benefits of commuter rail in eastern Massachusetts using spatial hedonic price functions. *Transportation* **2006**, *33*, 21–43. [CrossRef]
32. Yang, L.; Chen, Y.; Xu, N.; Zhao, R.; Chau, K.W.; Hong, S. Place-varying Impacts of Urban Rail Transit On Property Prices In Shenzhen, China: Insights for value capture. *Sustain. Cities Soc.* **2020**, *58*, 102140. [CrossRef]
33. Matthews, J.W.; Turnbull, G.K. Neighborhood Street Layout and Property Value: The Interaction of Accessibility and Land Use Mix. *J. Real Estate Financ.* **2007**, *35*, 111–141. [CrossRef]
34. Xiao, Y.; Orford, S.; Webster, C.J. Urban Configuration, Accessibility, and Property Prices: A Case Study of Cardiff, Wales. *Environ. Plan. B Plan. Des.* **2016**, *43*, 108–129. [CrossRef]
35. Shen, Y.; Karimi, K. The Economic Value of Streets: Mix-scale Spatio-functional Interaction and Housing Price Patterns. *Appl. Geogr.* **2017**, *79*, 187–202. [CrossRef]
36. Xiao, Y.; Li, Z.; Song, X. Estimating the Value of Street Layout via Local Housing Market: A Empirical Study of Nanjing, China. *Urban Dev. Stud.* **2015**, *22*, 6–11. (In Chinese)
37. Yang, Y.; Li, T.; Feng, X. Analysis on the Pattern of the House Price and Its Driving Forces in the Dwelling Space in the Urban District of Xi'an. *Areal Res. Dev.* **2015**, *34*, 68–74. (In Chinese)
38. Gu, H.; Meng, X.; Shen, T.; Chen, H.; Xiao, F. A Study on the Influence of Urban Road Network on the Housing Price in Guangzhou Based on sDNA Model. *Mod. Urban Res.* **2018**, *6*, 2–8. (In Chinese)
39. Gu, H.; Lin, Z.; Shen, T.; Mengzhen, H.; Liu, Z. Study on the Highway Network of Urban Agglomeration in the Middle Reaches of the Yangtze River Based on Space Syntax Theory. *Areal Res. Dev.* **2018**, *37*, 24–29. (In Chinese)
40. Liu, X.; Zhu, T. Analysis on the Applicability of Hedonic Price Method in Real Estate Valuation. *Foreign Econ. Relat. Trade* **2020**, *0*, 61–63. (In Chinese)
41. Yang, L.; Zhou, J.; Shyr, O.F.; Huo, D.D. Does Bus Accessibility Affect Property Prices? *Cities* **2019**, *84*, 56–65. [CrossRef]

42. Cervero, R.; Duncan, M. Land value impacts of rail transit services in Los Angeles County. In *Report Prepared for National Association of Realtors*; Urban Land Institute: Washington, DC, USA, 2002.
43. Yang, L.; Chau, K.W.; Szeto, W.Y.; Cui, X.; Wang, X. Accessibility to transit, by transit, and property prices: Spatially varying relationships. *Transp. Res. Part D Transp. Environ.* **2020**, *85*, 102387. [CrossRef]
44. Yang, L.; Chu, X.; Gou, Z.; Yang, H.; Lu, Y.; Huang, W. Accessibility and proximity effects of bus rapid transit on housing prices: Heterogeneity across price quantiles and space. *J. Transp. Geogr.* **2020**, *88*, 102850. [CrossRef]
45. Wei, X.; Zhang, W.; Wang, C.; Xu, G. The Impact of Urban Rail Transit on Surrounding Residential Prices–Line 1 of Chengdu Metro as an Example. *Modern Appl. Sci.* **2012**, *6*, 58. (In Chinese) [CrossRef]
46. Zhou, Z.; Chen, H.; Zhang, A. The Effect of a Subway on House Prices: Evidence from Shanghai: The Effect of a Subway on House Prices. *Real Estate Econ.* **2019**, *49*, 199. (In Chinese) [CrossRef]
47. Hu, Z.; Legeza, V.P.; Dychka, I.A.; Legeza, D.V. Mathematical Modeling of the Process of Vibration Protection in a System with two-mass Damper Pendulum. *Int. J. Intell. Syst. Appl.* **2017**, *9*, 18–25. [CrossRef]
48. Ahmed, N.; Rabbi, S.; Rahman, T.; Mia, R.; Rahman, M. Traffic Sign Detection and Recognition Model Using Support Vector Machine and Histogram of Oriented Gradient. *Int. J. Inf. Technol. Comput. Sci.* **2021**, *13*, 61–73. [CrossRef]

sustainability

Article

Relationship of Residential Location Choice with Commute Travels and Socioeconomics in the Small Towns of South Asia: The Case of Hafizabad, Pakistan

Houshmand Masoumi [1,2,*], Atif Bilal Aslam [3], Irfan Ahmad Rana [4], Muhammad Ahmad [3] and Nida Naeem [3,4]

1 Center for Technology and Society, Technische Universität Berlin, 10623 Berlin, Germany
2 Department of Transport and Supply Chain Management, College of Business and Economics, University of Johannesburg, Johannesburg 2006, South Africa
3 Department of City and Regional Planning, University of Engineering and Technology, Lahore 54890, Pakistan; atif.aslam@uet.edu.pk (A.B.A.); 2018mscrp8@student.uet.edu.pk (M.A.); nida.urp19@student.nust.edu.pk (N.N.)
4 Department of Urban and Regional Planning, School of Civil and Environmental Engineering, National University of Sciences and Technology, Islamabad 44000, Pakistan; iarana@nit.nust.edu.pk
* Correspondence: masoumi@ztg.tu-berlin.de

Abstract: The existing literature of emerging markets fails to provide evidence to clarify if people choose their residential location based on commuting to work or other socioeconomic or household factors. The present paper seeks to provide such evidence in South Asia using the case study of a small city in Pakistan. This exploratory study was facilitated by primary data collected from 365 adults in Hafizabad, Pakistan, using face-to-face interviews in 2018. Two research questions were answered: (1) with what socioeconomic or mobility-related variables are the residential self-selections correlated? (2) how strong is the possible association of commuting to work to residential location choices compared to other factors, including social, economic, and family-related issues? The results of Chi-square tests and Proportional Reduction in Error analyses show that the three variables of neighborhood place, gender, and housing tenure type are associated with residential location choices. These findings are partly in line with studies on high-income countries, but gender and housing tenure are more specific to developing countries. Moreover, results of a Binary Logistic model show that marital status and house ownership of other household members define whether people choose their living place based on commuting rather than other socioeconomic and household issues. The finding of the latter variable contrasts with behaviors in high-income countries, whereas the former variable has some similarities. These findings highlight some contextual differences between house location selection in South Asia and other regions.

Keywords: residential location choice; urban transportation planning; commuting; housing; Pakistan

check for updates

Citation: Masoumi, H.; Aslam, A.B.; Rana, I.A.; Ahmad, M.; Naeem, N. Relationship of Residential Location Choice with Commute Travels and Socioeconomics in the Small Towns of South Asia: The Case of Hafizabad, Pakistan. *Sustainability* **2022**, *14*, 3163. https://doi.org/10.3390/su14063163

Academic Editors: Linchuan Yang, Yuanyuan Guo, Wenxiang Li, Yaoming Zhou and Jixiang Liu

Received: 13 January 2022
Accepted: 4 March 2022
Published: 8 March 2022

1. Introduction

The correlations between residential location choices of the inhabitants of urban areas in high-income countries with different behaviors such as urban mobility choices are important for urban transportation researchers because they can influence the correlations between urban travel behaviors and the built environment. In other words, if residential self-selections meaningfully affect the mobility behaviors such as mode choice and travel distances, including commuting lengths, then it would be difficult to claim that the built environment can influence mobility behaviors and decisions. This may be true, particularly in relation to selecting residential places near the workplace to shorten the commuting distance. It is possible to hypothesize the commuting preferences and attributes of such people. Therefore, it is important to understand residential location choices which define their commuting characteristics.

The relationship between residential self-selection and travel behavior is complex, and the built environment plays a significant role in determining it [1–3]. Other important determinants that explain the relationship between residential choices and travel behavior are life choices in other relevant domains such as health and environment [4,5], decision-making arrangement at the household level, and commuting distance [6]. Cao and Yang (2017) found that the built environment has a significant effect on the commuting patterns even after controlling the effect of the residential self-selection [7]. However, the above evidence on the correlations between residential self-selections and commuting is mostly related to high-income countries. A very large part of the world's regions, including emerging markets and developing countries, represent a small proportion of the evidence. Due to the close relationship between mobility behaviors and decisions, on the one hand, and culture and climate, on the other hand, it can be hypothesized that context can have an undeniable role; however, because there is limited empirical evidence, it often cannot be claimed that several issues in urban planning and mobility planning can only be concluded based on evidence related to high-income countries. As a result, such conclusions are not valid to be the basis of mobility planning in emerging markets.

The present paper aims to understand the relationship between socioeconomic conditions and mobility patterns with residential location choices and preferences. It hypothesizes that the residential location choices in the developing countries are less affected by the commuting to work pattern as compared to the developed countries. Moreover, the correlates of some of the household-related variables such as household size are different in developing countries as compared to the same factors in high-income countries. These hypothetical differences have roots in the cultural differences and lifestyles of people in South Asia. To test these hypotheses, the small cities of the South Asian region are focused, exemplified by the city of Hafizabad, Pakistan, as a case study.

The paper continues with a short literature review on the correlates of residential self-selection. Then, the methods applied for testing the hypotheses of this study and its case study area, Hafizabad, Pakistan, are introduced. Then, the findings of the general correlations of residential location choices and different socioeconomic and mobility-related factors are presented. Finally, the findings of the South Asian context are compared with those of the existing literature, the majority of which come from high-income countries.

2. Correlates of Residential Location Choices

Travel behavior studies recognize residential location choices or self-selection as an integral part of understanding land use and transportation interactions [8]. It is being extensively used for launching relevant policy interventions for a sustainable transport system [9]. The public choices in choosing a residential location are primarily based on their travel options and priorities [10]. Numerous variables have been used in past studies to understand the correlations and determinants of residential location choices. Schirmer et al. (2014) classified these location variables within the categories of the built environment, socioeconomic environment, points of interest, and accessibility [3]. Frenkel et al. (2013) found socioeconomics, commuting time, and housing affordability as the primary factors of the residential location choice [11]. Orvin and Fatmi (2021) identified life-cycle events, accessibility, and socio-demographics as the key factors in determining the residential location choice [12]. Morency and Verreault (2020) found that a well-considered residential location choice can considerably reduce the commuting distances and as a result could also cause increasing walking, cycling, and public transport trips [13]. In some of the studies conducted in the developed world, social interactions and neighbors with similar socioeconomic backgrounds have also been a significant factor in residential location choices [14]. Other important determinants of residential location choices have been reported as the quality of schools [15], accessibility to services and jobs [16,17], mobility attitudes, the built environment [18–22], and the affordability and neighborhood characteristics [23,24].

Although numerous studies have been done in developed countries, travel behavior varies among different populations and regions due to socioeconomic conditions, housing types, norms, and attitudes [25]. Therefore, it is imperative to understand residential location choices for developing countries, specifically the South Asian region. There are limited studies on residential choices and self-selection in developing countries. Masoumi (2019) observed that residential location choices play a vital role in mode choice selection in Tehran, Istanbul, and Cairo [26]. In one of his recent studies for the same case study areas, Masoumi (2021) identified neighborhood characteristics, accessibility, commuting distance, public transit trips, and individual characteristics such as age and driving license that affect the residential location choice [27]. Masoumi (2013) also found a significant role of socioeconomics in determining the residential location choices in Tehran, Iran [28]. In another study conducted in Alexandria, Masoumi et al. (2021) identified neighborhood characteristics, availability of transportation modes, and affordability as the strongest determinants of the residential location choices [29]. Ibrahim (2017) also found the availability of transportation modes as the leading determinant of the residential location choices in Alexandria [30]. Albayrak et al. (2019) argued that housing affordability and travel behavior shapes the housing choices of the residents of the mono-centered city. In contrast, the situation in a poly-centered city like Istanbul is complex. Several factors such as individual preferences, job location, accessibility, and sociocultural factors determine housing location choices [31]. Salihoglu and Turkoglu (2019) also highlighted various factors such as housing and neighborhood characteristics, accessibility, and residential satisfaction that affect residential location preferences in Istanbul [32]. Ghazali et al. (2020) studied residential location choices in the city of Elmina, Malaysia, by conceiving a broader frame of the migration-related push-pull-mooring model. The study concluded that pull factors such as affordability and socioeconomic factors are responsible for residential location choices at the destination places. Certain push factors, such as the origin place, dissatisfaction, and high housing costs, also play a significant role in residential location choices [33]. Jiang and Zhang (2021) found that neighborhood characteristics, housing price, accessibility to transportation, and entertainment places are important determinants of location choices for housing purchase in Anyue County, China [34]. Aung and Vichiensan (2019) identified housing characteristics, neighborhood quality, commuting time, and ethnicity as significant factors affecting the residential location preferences in Myanmar [35]. Many other studies in the developing world have found similar determinants of residential location choices such as accessibility and travel behavior [7,36], neighborhood and socioeconomic characteristics [37], affordability and security [38], convenience and comfort [39], and religious factors [40].

A study by Munshi (2016) observed that residential location choice is important to be considered for determining mode choice in Rajkot, India [41]. Pandya and Maind (2017) found distance to the central business district, housing affordability, and family income to be significant factors that affect residential location choice in the Mumbai Metropolitan Region [42]. Aslam et al. (2019) conducted a study on a similar topic in the same small city of Hafizabad, Pakistan, and, through descriptive analysis, found affordability and availability of utility services to be the leading factors of residential location choices [43]. De and Vupru (2017) found socioeconomics, accessibility to the workplace, and amenity facilities to be important factors in determining housing location choice and the rental values of the residents of a small city of Dimapur Town in Nagaland, India [44]. Digambar et al. (2010) found housing ownership and housing type to be significant factors affecting the residential location choices of high- and middle-income groups in Nagpur, India [45]. Rehman and Jamil (2021) reported commuting cost and housing rent to be the determinants of residential location choice in the twin cities of Rawalpindi and Islamabad [46]. Some other studies have also revealed the importance of socioeconomics in shaping housing location choices in the South Asian region [47]. For example, Choudhury and Ayaz (2015) found the quality of educational institutions and house rents as the leading determinants of residential location choices in Bangladesh [48]. Shawal and Ferdous (2014) did a similar study with workers of garment factories in Dhaka, Bangladesh. They found a range of factors, including socioeconomics, affordability, accessibility to services, and commuting distance, which affected residential location choices [49]. Thus, it is imperative to understand the residential location choices in other South Asian cities for improving land-use transportation dynamics.

3. Materials and Methods

Based on the literature review and the knowledge gaps, the current study seeks to answer the following research questions: (1) with what socioeconomic or mobility-related variables are the residential self-selections correlated? (2) how strong is the possible association of commuting to work to residential location choices compared to other factors, including social, economic, and family-related issues? This study hypothesizes that unlike some studies conducted in Western countries, residential location choices in South Asian countries are less influenced by commuting to work. Thus, it is easier to study the correlations between urban travel behaviors and the built environment in that context. This is because if the hypothesis is tested to be true, residential location choices in the South Asian context would work more as a constant than a variable to cause changes in other domains, most importantly, the travel behavior and the characteristics of the built environment.

A small city of Hafizabad located in the upper central Punjab region of Pakistan was chosen to conduct this study as the monocentric character of the city offered some advantages for reliably concluding this study with a smaller sample size. The population of Hafizabad, according to the 2017 Census, has been reported as 245,784. Furthermore, there were 37,270 housing units in Hafizabad with a household size of 6.6 persons—slightly higher than the national average of 6.5 [50]. Despite being a small city, it is well connected with other urban places in the surroundings. Gujranwala, the fifth largest city in Pakistan with a population of 2.03 million [50], is located only around 55 km away in the East, enabling traveling between these two cities [51]. The urban fabric of the city consisted of many layers dating back to the Mughal dynasty, followed by the British empire, which exercised Victorian architecture during the colonial times. Since the independence of Pakistan in 1947, post-partitioned time urban layers have also been added to the urban landscape of Hafizabad.

This study is based on a survey undertaken in Hafizabad in 2018, which led to a validated sample of 365 residents. Cochran's (1963) formula was applied in determining the sample size and confidence interval for conducting this study. A two-stage sampling technique was used where in the first stage, four neighborhoods were selected based on their distinct land use and built environment characteristics. The calculated sample was equally distributed among the selected neighborhoods. In the second stage, a probabilistic random sampling technique was used to complete the sample size for ensuring its representativeness for the overall population. The random sampling offered an opportunity to handle the cases of refusals as the field surveyors moved to the next respondents in all such cases [52]. The survey method was face-to-face interviews in the four case study neighborhoods of the study, which contained a combined population of 19,042 inhabitants. The survey resulted in individual and household response rates of 1.92% and 12.65% and confidence levels of ±5.08% and ±4.79% for individual and household questions. The overall data collection was performed as an exploratory survey. The response rates and confidence levels of each neighborhood have been summarized in Table 1. The full details of the data collection have already been published by Aslam et al. (2019) [43].

The most important factors in connection with residential self-selections were selected to be applied in statistical analyses. The neighborhood was applied to the tests because it is an indirect index of socioeconomic status. As an example of the difference in the economic levels of the neighborhoods, house prices can be raised. The cheapest houses are found in Hassan Town (29%), whereas the most expensive houses are in Nawab Colony (16%). Personal variables include age, gender, marital status, and employment. For cultural reasons, gender was only designated as two categories, making up a dummy variable. Household variables include vehicle ownership (bike and car), type of housing, previous relocation, time of relocation, house ownership of other members, the present status of housing, and the actual price of a house. Finally, two variables represent mobility habits: travel time and mode choice. All of these data were designed as categorical (and binary) variables, the frequencies shown in Table 2.

Table 1. The survey characteristics [43].

Neighborhoods	Projected Population	Number of Households	Number of Interviewed Subjects	Neighborhood-Level Validated Sample Size (n)	Response Ratio for Individual Variables (%)	Response Ratio for Household Variables (%)	Confidence Interval for Individual Variables	Confidence Interval for Household Variables (%)
Muhallah Hassan Town	7.861	1.191	100	100	1.27	8.40	9.74	9.38
Muhallah Shareef Pura	3.298	500	100	100	3.03	20.00	9.65	8.77
Gali Haji Miraaj Din	3.584	543	100	100	2.79	18.42	9.66	8.86
Nawab Colony	4.299	651	98	65	1.51	9.98	12.06	11.54
Total	19.042	2.885	398	365	1.92	12.65	5.08	4.79

Table 2. The frequencies of independent and explanatory variables of this study.

Category	Sub-Category	Frequency	Percent	Category	Sub-Category	Frequency	Percent
Residential Location Choice	Commuting	61	16.7	**Age**	Between 21–35	128	35.1
	Other Factors	304	83.3		Between 36–45	164	44.9
					46 and above	72	19.7
Neighborhood (Socioeconomic Status)	Gali Haji Miraaj Din	100	27.4				
	Shareef Pura	100	27.4				
	Hassan Town	100	27.4				
	Nawab Colony	65	17.8	**Gender**	Male	308	84.4
					Female	57	15.6
Marital Status	Single	44	12.1				
	Engaged	14	3.8	**Car and Bike Ownership**	No car	83	22.7
	Married	299	81.9		One car	17	4.7
	Widow	7	1.9		No bike	50	13.7
					One bike	205	56.2
					Two bikes	7	1.9
Employment	Full-time job	286	78.4		More than two bikes	3	0.8
	Part-time job	40	11.0				
	Work at home	24	6.6	**Other Types of Housing (tenure)**	Owned	321	87.9
	Searching for a job	8	2.2		Rent	44	12.1
	Retired	7	1.9				

Table 2. *Cont.*

Category	Sub-Category	Frequency	Percent
Type of House	Private apartment	1	0.3
	Self-built house	317	86.8
	Others	47	12.9
Residential Location Choice-Categorical	I afford this house	52	14.2
	Proximity to work	49	13.4
	Family's asset	149	40.8
	Availability of transportation	12	3.3
	Nice neighborhood	57	15.6
	Proximity to family/relatives	35	9.6
	Nearby downtown	8	2.2
	Social standing of the area	3	0.8
Previous House Relocation	Yes	126	34.5
	No	239	65.5
Last Relocation Time	Less than 2 years	34	9.3
	From 2 to 10 Years	54	14.8
	More than 10 Years	40	11.0
Present Status of Housing Units owned by other members	MissingVacantOccupied	300 34 31	82.2 9.3 8.5

Category	Sub-Category	Frequency	Percent
Travel Time	Less than 30 Min	260	71.2
	Between 30–60 min	51	14.0
	Between 61–90 min	3	0.8
	90 min, or more	3	0.8
	I work outside Hafizabad	39	10.7
Mode of Transportation	Missing	23	6.3
	Walking	147	40.3
	Private car	9	2.5
	Bus or minibus	19	5.2
	Train or tram	3	0.8
	Taxi	5	1.4
	Ride in a friend's Car	9	2.5
	Transport facility offered by the company	4	1.1
	Motorcycle	146	40.0
House Ownership of Other Household Members	1	56	15.3
	2	7	1.9
	4	2	0.5
Actual Price of House	Less than PKR * 1.5 million	55	15.1
	Between PKR 1.5–3.0 million	180	49.3
	Between PKR 3.0–4.5 million	59	16.2
	PKR 4.5 million and above	31	8.5

* 1 PKR = 0.0057 USD (Source: https://www.forex.pk/currency-converter.php dated 7 February 2022).

To answer the first research question of this study, the residential location choice variable produced by the questionnaire was applied in categorical form. The choices included eight categories of "I afford this house", "proximity to work", "family's assets", "availability of transportation", "nice neighborhood", "proximity to family/relatives", "nearby downtown", and "social standing of the area". These options were the results of open-ended questions in previous pilot studies in other developing countries. The correlations between this variable and ten other household, socioeconomic, spatial, and mobility-related variables were tested by the Chi-square test of independence on a univariate basis. The null hypothesis was that there was no association between the two variables. This hypothesis was rejected by a *p*-value of less than 0.05. Since the Chi-square test of independence does not show the strength of associations, the Proportional Reduction in Error (PRE) method was applied to analyze the strength of associations. The variables of this analysis were all nominal, so the Cramer's V measure was estimated for the tests between each pair of variables according to Formula (1).

$$V = \sqrt{\chi^2/(N)(\min r - 1, c - 1)} \quad (1)$$

where χ^2 is the value of Chi-square measure, N is the number of the subjects in the sample, and (min $r-1$, $c-1$) is the minimum value of either the number of residential location choice $-$ 1 or the number of the other variable $-$ 1. When χ^2 produced insignificant values ($p > 0.05$), it was concluded that the two variables were associated; thus, their PRE strength was estimated by the Cramer's V using the following thresholds:

$0 < V < 0.10$: weak association between variables; $0.1 < V < 0.30$: moderate association between variables; and $V > 0.30$: strong association between the variables.

To answer the second research question, the categorical variable of residential location choice was transformed into a binary variable with reasons related to commuting to work, including proximity to work and availability of transportation infrastructure versus other factors including affordability, family's asset, nice neighborhood, proximity to family/relatives, near to downtown, and social standing of the neighborhood. This dummy variable is assumed to represent choosing the living location homogenously in response to commuting needs, e.g., the mode of commuting does not influence the residential self-selection. Then, the dummy variable was taken as the dependent variable of a Binary Logistic (BL) model, and the potentially most effective variables were applied as explanatory variables. Modeling was repeated nine times, during which transport mode, the average area of a house, type of house, gender, age, vehicle ownership, employment, and travel time were eliminated from the model, respectively.

4. Results

According to the approach explained in the methodology section, the categorical variable of residential location choice (including eight choices) was tested against ten variables (neighborhood, age, gender, marital status, employment, car and bike ownership, type of house, travel time, mode of transportation, and house ownership of other household members). Where an association is found, the Cramer's V value indicates the strength. The results of these univariate hypothesis testing are listed in Table 3. The *p*-values of three variables of "neighborhood", "gender", and "type of house (tenure)" show highly significant or significant association with a residential location choice. This indicates that knowing where the people live, their gender, or what type of tenure they have can help us predict the reasons behind their house location choice. The value of Cramer's V shows the strength of these correlations; however, these results do not reflect the pattern of the associated variables. The distribution patterns of residential self-selections versus the three variables can be found in Figure 1, equivalent to contingency tables of each pair of variables.

Table 3. The Chi-square test results and Proportional Reduction in Error (PRE) analysis for residential location choice with socioeconomics and mobility variables in Hafizabad.

Variable	Pearson Chi-Square Value	Df	*p*-Value	Cramer's V
Neighborhood	50.73	21	<0.001	0.215
Age	17.49	14	0.231	0.155
Gender	14.49	7	0.043	0.199
Marital status	22.71	21	0.359	0.144
Employment	25.06	28	0.625	0.131
Car and bike ownership	30.05	35	0.706	0.128
Type of house	54. 04	14	<0.001	0.272
Travel time	30.96	28	0.319	0.147
Mode of transportation	53.15	56	0.583	0.144
House ownership of other household members	10.89	7	0.143	0.173

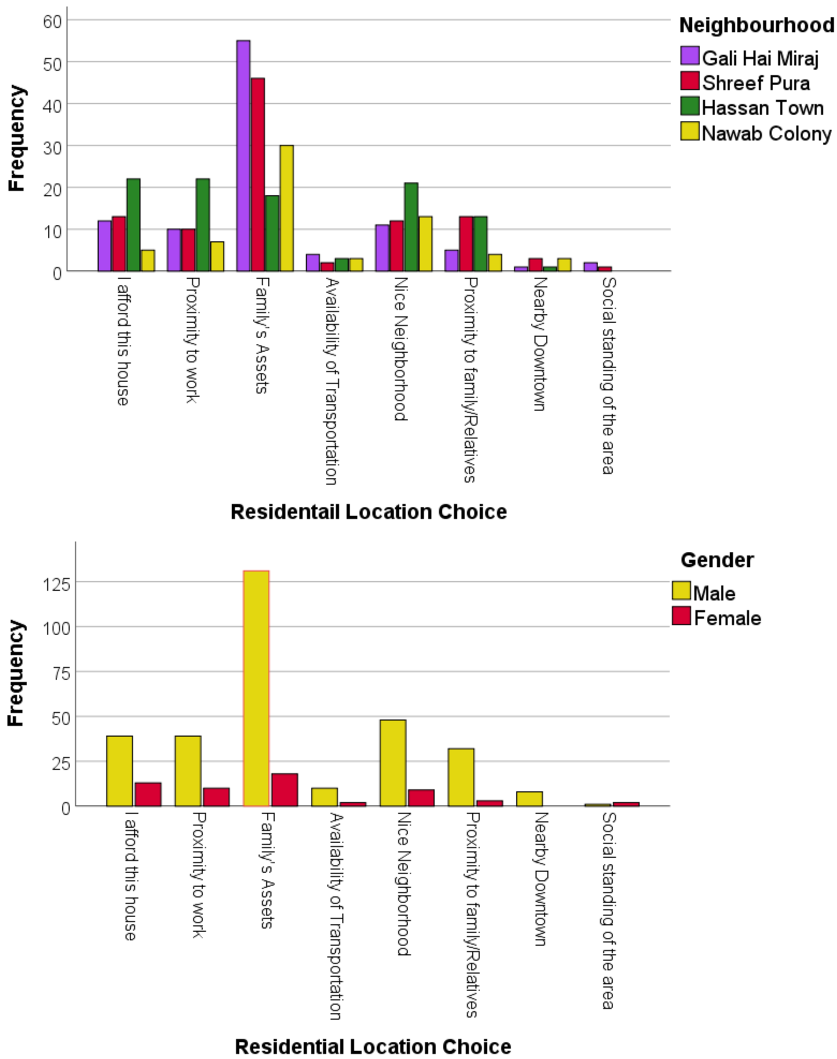


Figure 1. The pattern of frequencies of residential location choice and two correlated variables in Hafizabad, Pakistan.

According to these findings, the null hypothesis of no association between house location choices and the location of the neighborhoods of the respondents has been rejected ($p < 0.001$). Residential location choices are highly significantly associated with the neighborhood in which the respondents live. In other words, the distribution pattern of the responses about house location choices is in accordance with the frequency of the district of the city in which the respondents live. This strength of this association is moderate (Cramer's V = 0.215). The frequencies of location choices in different neighborhoods (Figure 1) show that the reasons follow almost the same pattern in all areas. A relative exception is Hassan Town, a full-grid newly developed area (post-2000) compared to other selected neighborhoods of this study. However, the small deviances of Hassan Town seen in Figure 1 have not had any significant effect on the results of the hypothesis testing.

Gender is a variable that was proved by the findings to have a significant association with residential location choice ($p = 0.043$). The strength of this correlation is also moderate (Cramer's V = 0.199). Figure 1 shows how males and females have similar patterns of residential self-selection. This is understandable through the cultural atmosphere and lifestyles in a small town in Pakistan, where household members live together, and the decisions are taken centrally by senior members. This contrasts with individual decisions made by small household units in Western societies.

Finally, type of tenure is the last associated variable with residential locations choices. This association is also highly significant ($p < 0.001$) and has a moderate strength (Cramer's V = 0.272) (Table 3). The distribution patterns indicate that whether the houses are owned or rented, the residential self-selections have the same pattern and distribution. In other words, the location choices are not under the influence of renting or buying the living unit. People's way of thinking is the same regarding choosing their living location when they plan to buy or rent a house.

The final form of the BL model can help us investigate the relationship between residential self-selection and commuting to work. Two variables of marital status and house ownership of other household members are associated with deciding to buy or rent a house because of making commuting easier against other reasons for location choice (Table 4).

Table 4. Binary logistic model for residential location choices in Hafizabad (commuting vs. other factors).

Variable	B	S.E.	Wald	df	p	β
Marital Status = Single	3.446	1.084	10.100	1	0.001	31.379
Marital Status = Engaged	2.465	1.142	4.659	1	0.031	11.767
Marital status = Married	1.479	0.449	10.850	1	0.001	4.388
House ownership of other household members = 1	−0.743	0.386	3.703	1	0.054	0.476
Actual price of house in real estate market	0.156	0.190	0.670	1	0.413	1.169

The first important variable, all categories significantly associated with choosing house location based on commuting, is marital status. The reference category of this variable is "widow", which per se is not an important category, but taking this category as a reference helps understand the relationship between other categories, particularly single and married people. Single and married groups are highly significant in the model ($p = 0.001$). The proportion of the β values of the two categories indicates that single respondents of the sample are 7.15 times more likely to choose their house place based on commuting to work rather than other factors like socioeconomics, etc. Similarly, single people are 2.66 times more likely to select their house location in ease of commuting compared to engaged people. Finally, engaged people are 2.68 times more likely than married people. These findings emphasize the importance and effectiveness of life course events like marriage, having children, etc., in behaviors related to mobility.

The other significant factor in the model is related to house ownership by other household members. People living in a household in which one of the other members, like siblings, have one house are 48% more likely to choose their house location based on commuting than people living in families in which other members have four houses.

Although this relationship is marginally significant (p = 0.054), it shows that having more available houses directly by the household members may be in relation to the selection of new houses by other family members based on commuting or other factors. This relation seems to be linked with complex cultural issues found in contexts with more traditional central families, where the household sizes are large, and there is a close relationship between household members.

Finally, the last variable in the model, the actual price of a house in the real estate market, is not significant (p = 0.413). Still, it was kept in the model to increase the validity results of the model, and at the same time, it functions as a control variable. In other words, when the house price is controlled for (it is fixed), the marital status and house ownership of other household members will be significant in relation to choosing a house location based on commuting versus other issues.

The results of the model validity tests can be seen in Table 5. The Nagelkerke R^2 equals 0.604, which means the model can predict 60.4% of the variance of house location choices chosen from commuting to work against other reasons. Although this is not a bad R^2, more complex and detailed data can help get better results in future models. The results of the Omnibus Test (p < 0.001) and Hosmer and Lemeshow Test show that the model is valid and performs well (Table 5).

Table 5. The model validation test results of the binary logistic model of residential location choice in Hafizabad.

Model Summary	**−2 Log Likelihood** 253.82	**Cox & Snell R^2** 0.453	**Nagelkerke R^2** 0.604
Omnibus Tests of Model Coefficients	Chi-square 195.33	Df 5	p <0.001
Hosmer and Lemeshow Test	Chi-square 7.83	Df 5	p 0.165

5. Discussion

The results identified three main variables, i.e., neighborhood, gender, and housing tenure, as the correlates of residential self-selection in Hafizabad. In contrast, all other investigated variables did not show any significant association. This shows the importance of social life maintained through various networks at the neighborhood level, neighborhood characteristics, demographic structuring of households with respect to decision-making arrangements, and owner-occupied housing tenures for residential location choices of the people residing in smaller cities of the South Asia region. These findings have some similarities and differences with the results emerging from the developed countries. The variables of neighborhood-related attributes such as nice neighborhood [3,18,24], presence of family, social contacts, and people having similar socioeconomic status [14], etc., have also been reported as a significant determinant of residential location choices in the developed world. However, the role of gender and housing tenure in deciding the residential self-selection in developing countries is different than the results surfacing from the studies conducted in the developed world.

The majority of the respondents (84.4%) of this study were male, and the Chi-square test results imply that they were the main actors choosing the residential locations for their families. This reflects the dominantly patriarchal character of Pakistani society, where male heads of the households make the key decisions [53]. A study found gender to be a significant factor of residential location choices in Alexandria, Egypt [29]. In Nigeria, a study found low involvement of women in deciding residential locations [31]. Generally, these findings are context-specific to the developing countries, and finding comparable results for the developed countries is difficult. A rare study conducted in a high-income country found significant gender-based differences in choosing the residential locations in Tel Aviv [54]. In many other studies conducted in the developed world, both males and females decide together for their residential locations [55]. Housing tenure was also found to be another significant variable that determines the residential location choices of the residents of Hafizabad, Pakistan, which is mainly shaped up by the majority type of owner-occupied housing tenure (87.9%). In the absence of sufficient social safety nets for the masses, owning a house is generally an asset accumulation strategy of the people, which gives them a feeling of security and protection in times of need. Another study also found owner-occupation of housing units in the shape of family assets as determinants of residential location choices in Alexandria, Egypt [29]. However, this finding has not surfaced as such for the cases of developed countries where the ratio of owner-occupied housing units was found less than many developing countries.

The life events such as marriages, having children, and relocation to new places also have significant impacts on mobility behaviors. The tested model in this study provided evidence of a strong association of marital status with the residential location choices based on commuting patterns. This finding conforms with some of the results coming from the developed world. A study identified marital status to significantly affect the residential relocation and associated travel behavior in the Metro Vancouver region [56]. However, there are studies from the developed world that do not provide evidence for similar findings. Some studies identified residential relocation and related changes in the built environment to affect travel patterns and car ownership in Cologne significantly [17,21]. Researchers also found similar findings for German cities; however, they argued that residential relocation affects travel behavior differently across varying scales [57,58]. Such studies did not mention the important life event of marriage as an affecting factor for residential location choices based on travel behavior. Another variable that emerged as having a significant association with residential location choices based on commuting behavior was house ownership by other household members. This is an interesting context-specific finding which has not been reported in the studies conducted in the developed world. A joint family system and larger household sizes in the developing world could be the reasons which result in having more than one household living in one house in many cases. As per the last Census activity in Pakistan, the average household size is 6.39 [50] higher than the household sizes in many developed countries. However, such higher household sizes and joint living arrangements encourage some household members to own other house/s for possible residential relocation in future times due to larger individual household size. The only evidence available from the studies done in the developed world is the differentiation between the individual preferences and joint decision making within households for residential location choices such as [59].

These findings would have a significant bearing on urban planning practices if integrated with the policy formulation to efficiently plan and manage the urban places along with the sustainability principles, particularly in developing countries. The neighborhood unit is an important scale that significantly affects the residential location choices of the residents of the urban places. However, many urban planning instruments operate on a city-scale without understanding varying dynamics across different neighborhoods. This necessitates integrating neighborhoods as an important unit of analysis while devising urban planning responses and service delivery mechanisms for addressing the urban housing issues. Additionally, to promote an inclusive urban planning paradigm, many urban planning initiatives for housing delivery need to be based on gender-sensitive planning. Owner-occupied housing tenure has also been found out an important determinant of residential location choices as the family asset is generally kept on transferring to the next generations. A policy response may be devised to discourage the change of land uses of such owner-occupied housing units typically located within the central parts of the city to make them available for residential activity. This will be an important planning intervention to keep the job–housing balance right within the central parts to promote the balanced growth of the urban places. This planning intervention has also been extensively applied in the central parts of Freiburg, Germany, to serve the same purpose as stated above [60]. The life events of the peoples' biography are also essential to understand their residential location choices and associated travel behavior. This study reports marriage leads to an increased number of households and larger family sizes within the same housing units as a significant predictor of peoples' residential location choices based on commuting behavior. This warrants the consideration of marriages and the increased number of households as the basis for estimating housing shortage and demand forecasting, which will shape up the required urban housing policy response in a meaningful way. This finding is also essential for urban transport planners. If taken into account, this may result in an effective transport planning response aiming at shorter commuting distances and promoting active modes of travel.

As the study premise was a small city with a mono-centric character located in a developing country of South Asia, the findings of this study are very much context-specific. This means generalizing to portray the situation of larger cities of the same region based on these findings would not be possible as larger cities with poly-centric characteristics are much more complex. Additionally, the data was collected pre-COVID-19 times, so the results may not be an accurate depiction of the situation of current post-COVID-19 times as the COVID-19 pandemic has largely affected the transportation and mobility behavior of the people across the globe [61–64].

6. Conclusions

This study mainly addresses the two key research questions about residential location choices in relation to commuting behavior and other socioeconomic factors in the small city of Hafizabad, Pakistan. The study concludes that the factors of the neighborhood, gender, and housing tenure are associated with the residential location choices of the people. Apart from the variable of the neighborhood, other reported variables are not like those which are reported in the developed world. Additionally, two significant variables of marital status and house ownership of other household members have been surfaced as associated with the residential location choices based on commuting behavior. This finding draws lesser similarities and more differences with the studies conducted in the developed world. Overall, the study concludes that the residential location choices of the people in relation to the commute travel and socioeconomics in the developing countries have lesser similarities and more differences with the people behaviors in these domains in the developed countries. These findings bear important implications for urban planning interventions by relevant actors to promote sustainable urban development and mobility.

Similar studies also need to be done in the larger cities of the Global South. Considering the complexities and the scale of the problems faced by the larger urban places in the developing countries, such studies will be even more important from an effective policy formulation point of view. However, such studies need to incorporate the COVID-19 factor while designing the research methodology, as it is still an ongoing phenomenon. Some studies might also be needed to report on the mental mapping to understand the peoples' perception with respect to their residential location preferences. Findings of such studies may help the relevant actors formulate a policy response aiming to steer the revealed preferences towards more favorable and sustainable housing supply solutions.

Author Contributions: Conceptualization, H.M. and A.B.A.; methodology, H.M.; software, H.M.; validation, H.M.; formal analysis, H.M.; investigation, H.M., A.B.A. and I.A.R.; data curation, A.B.A., M.A. and N.N.; writing—original draft preparation, H.M., A.B.A., I.A.R., M.A. and N.N.; writing—review and editing, H.M.; visualization, H.M.; supervision, H.M. and A.B.A.; project administration, H.M. and A.B.A.; funding acquisition, H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This study was undertaken with the support of the German Research Foundation (DGF) as the research project "Urban Travel Behavior in Large Cities of MENA Region" (UTB-MENA) with the project number MA6412/3-1.

Institutional Review Board Statement: Ethical review and approval were waived for this study, as no interventions were included in the data collection about humans.

Informed Consent Statement: Written informed consent has not been obtained from the respondents of the interviews of this study, due to the cultural conditions, in which data were collected.

Data Availability Statement: At the time of the publication of this paper, the data have not been shared on a repository.

Acknowledgments: We acknowledge support from the German Research Foundation and the Open Access Publication Fund of Technische Universität Berlin.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Guan, X.; Wang, D.; Jason Cao, X. The role of residential self-selection in land use-travel research: A review of recent findings. *Transp. Rev.* **2020**, *40*, 267–287. [CrossRef]
2. Lin, T.; Wang, D.; Guan, X. The built environment, travel attitude, and travel behavior: Residential self-selection or residential determination? *J. Transp. Geogr.* **2017**, *65*, 111–122. [CrossRef]
3. Schirmer, P.M.; Van Eggermond, M.A.; Axhausen, K.W. The role of location in residential location choice models: A review of literature. *J. Transp. Land Use* **2014**, *7*, 3–21. [CrossRef]
4. Boone-Heinonen, J.; Gordon-Larsen, P.; Guilkey, D.K.; Jacobs, D.R., Jr.; Popkin, B.M. Environment and physical activity dynamics: The role of residential self-selection. *Psychol. Sport Exerc.* **2011**, *12*, 54–60. [CrossRef] [PubMed]
5. Zhang, J. Revisiting residential self-selection issues: A life-oriented approach. *J. Transp. Land Use* **2014**, *7*, 29–45. [CrossRef]
6. Guan, X.; Wang, D. Residential self-selection in the built environment-travel behavior connection: Whose self-selection? *Transp. Res. Part D Transp. Environ.* **2019**, *67*, 16–32. [CrossRef]
7. Cao, X.; Yang, W. Examining the effects of the built environment and residential self-selection on commuting trips and the related CO2 emissions: An empirical study in Guangzhou, China. *Transp. Res. Part D Transp. Environ.* **2017**, *52*, 480–494. [CrossRef]
8. Næss, P. Residential Self-Selection and Appropriate Control Variables in Land Use: Travel Studies. *Transp. Rev.* **2009**, *29*, 293–324. [CrossRef]
9. De Abreu e Silva, J. Spatial self-selection in land-use–travel behavior interactions: Accounting simultaneously for attitudes and socioeconomic characteristics. *J. Transp. Land Use* **2014**, *7*, 63. [CrossRef]
10. Mokhtarian, P.L.; Cao, X. Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transp. Res. Part B Methodol.* **2008**, *42*, 204–228. [CrossRef]
11. Frenkel, A.; Bendit, E.; Kaplan, S. Residential location choice of knowledge-workers: The role of amenities, workplace and lifestyle. *Cities* **2013**, *35*, 33–41. [CrossRef]
12. Orvin, M.M.; Fatmi, M.R. Modeling residential mobility decisions from a life history–oriented perspective. *Transp. Lett.* **2021**, 1–10. [CrossRef]
13. Morency, C.; Verreault, H. Assessing the Efficiency of Household Residential Location Choices. *Transp. Res. Rec.* **2020**, *2674*, 455–465. [CrossRef]
14. Kiefer, H. Residential Location Choice: The Role of a taste for similarity. *Int. J. Econ. Financ.* **2012**, *4*, 34–60. [CrossRef]
15. Guidon, S.; Wicki, M.; Bernauer, T.; Axhausen, K. The social aspect of residential location choice: On the trade-off between proximity to social contacts and commuting. *J. Transp. Geogr.* **2019**, *74*, 333–340. [CrossRef]
16. Baraklianos, I.; Bouzouina, L.; Bonnel, P.; Aissaoui, H. Does the accessibility measure influence the results of residential location choice modelling? *Transportation* **2018**, *8*, 11. [CrossRef]
17. Hu, L.; Wang, L. Housing location choices of the poor: Does access to jobs matter? *Hous. Stud.* **2019**, *34*, 1721–1745. [CrossRef]
18. Cao, X. Residential self-selection in the relationships between the built environment and travel behavior. Introduction to the special issue. *J. Transp. Land Use* **2014**, *7*, 1–3. [CrossRef]
19. Ettema, D.; Nieuwenhuis, R. Residential self-selection and travel behaviour: What are the effects of attitudes, reasons for location choice and the built environment? *J. Transp. Geogr.* **2017**, *59*, 146–155. [CrossRef]
20. Heinen, E.; Mackett, R.; van Wee, B.; Ogilvie, D.; Panter, J. Residential self-selection in quasi-experimental and natural experimental studies. *J. Transp. Land Use* **2018**, *11*, 939–959. [CrossRef]
21. Scheiner, J. Residential self-selection in travel behavior: Towards an integration into mobility biographies. *J. Transp. Land Use* **2014**, *7*, 15–29. [CrossRef]
22. Zang, P.; Lu, Y.; Ma, J.; Xie, B.; Wang, R.; Liu, Y. Disentangling residential self-selection from impacts of built environment characteristics on travel behaviors for older adults. *Soc. Sci. Med.* **2019**, *238*, 112515. [CrossRef] [PubMed]
23. Bayoh, I.; Irwin, E.G.; Haab, T. Determinants of residential location choice: How important are local public goods in attracting homeowners to central city locations? *J. Reg. Sci.* **2006**, *46*, 97–120. [CrossRef]
24. Lee, B.H.; Waddell, P. Residential mobility and location choice: A nested logit model with sampling of alternatives. *Transportation* **2010**, *37*, 587–601. [CrossRef]

25. Humphreys, J.; Ahern, A. Is travel based residential self-selection a significant influence in modal choice and household location decisions? *Transp. Policy* **2019**, *75*, 150–160. [CrossRef]
26. Masoumi, H.E. A discrete choice analysis of transport mode choice causality and perceived barriers of sustainable mobility in the MENA region. *Transp. Policy* **2019**, *79*, 37–53. [CrossRef]
27. Masoumi, H. Residential Location Choice in Istanbul, Tehran, and Cairo: The Importance of Commuting to Work. *Sustainability* **2021**, *13*, 5757. [CrossRef]
28. Masoumi, H.E. Residential self-selection and its effects on urban commute travels in Iranian cities compared to US, UK, and Germany. *Int. J. Soc. Sci.* **2013**, *7*, 877–881. [CrossRef]
29. Masoumi, H.; Ibrahim, M.R.; Aslam, A.B. The Relation between Residential Self-Selection and Urban Mobility in Middle Eastern Cities: The Case of Alexandria, Egypt. *Urban Forum* **2021**, *32*, 261–287. [CrossRef]
30. Ibrahim, M.R. How do people select their residential locations in Egypt? The case of Alexandria. *Cities* **2017**, *62*, 96–106. [CrossRef]
31. Albayrak, A.N.; Salihoğlu, T.; Salihoğlu, G. Qualified Labors' Residential Location Choice and Residential Satisfaction//Nitelikli İşgücünün Konut Alanı Yerseçimi Tercihi ve Memnuniyeti. *Megaron* **2019**, *14*, 122–132.
32. Salihoglu, T.; Türkoglu, H. Determinants of residential location choice in Istanbul: A longitudinal study. *Open House Int.* **2019**, *44*, 36–44. [CrossRef]
33. Ghazali, E.M.; Ngiam, E.Y.L.; Mutum, D.S. Elucidating the drivers of residential mobility and housing choice behaviour in a suburban township via push–pull–mooring framework. *J. Hous. Built Environ.* **2020**, *35*, 633–659. [CrossRef]
34. Jiang, J.; Zhang, J. Analysis of County Consumers' Housing Purchase Intention and Influencing Factors-Based on the Investigation of Anyue County. In Proceedings of the International Conference on Management Science and Engineering Management, Toledo, Spain, 1–4 August 2021; Springer: Cham, The Netherlands, 2021; pp. 820–834. [CrossRef]
35. Aung, T.H.H.; Vichiensan, V. A Stated Preference Experiment of Residential Location Choice in Mandalay. In Proceedings of the 2019 First International Conference on Smart Technology & Urban Development (STUD), Chiang Mai, Thailand, 13–14 December 2019; pp. 1–8. [CrossRef]
36. Wang, D.; Lin, T. Residential self-selection, built environment, and travel behavior in the Chinese context. *J. Transp. Land Use* **2014**, *7*, 5–14. [CrossRef]
37. Poku-Boansi, M.; Adarkwa, K.K. Determinants of residential location in the Adenta Municipality, Ghana. *GeoJournal* **2016**, *81*, 779–791. [CrossRef]
38. Ubani, P.; Alaci, D.S.; Udoo, V. Determinants of residential Neighborhood choice in a Nigerian Metropolis. *IOSR J. Humanit. Soc. Sci.* **2017**, *22*, 1–11. Available online: https://www.iosrjournals.org/iosr-jhss/papers/Vol.%2022%20Issue7/Version-11/A2207110111.pdf (accessed on 7 November 2021).
39. Wang, M.; Yang, Y.; Jin, S.; Gu, L.; Zhang, H. Social and cultural factors that influence residential location choice of urban senior citizens in China–The case of Chengdu city. *Habitat Int.* **2016**, *53*, 55–65. [CrossRef]
40. Sinniah, G.K.; Shah, M.Z.; Vigar, G.; Aditjandra, P.T. Residential location preferences: New perspective. *Transp. Res. Procedia* **2016**, *17*, 369–383. [CrossRef]
41. Munshi, T. Built environment and mode choice relationship for commute travel in the city of Rajkot, India. *Transp. Res. Part D Transp. Environ.* **2016**, *44*, 239–253. [CrossRef]
42. Pandya FMaind, S. Discrete choice model: Residential location choice. In ISFIRE Working Paper Series, 28. Available online: http://iire.in/ojs/index.php (accessed on 28 October 2021).
43. Aslam, A.B.; Masoumi, H.E.; Naeem, N.; Ahmad, M. Residential location choices and the role of mobility, socioeconomics, and land use in Hafizabad, Pakistan. *Urbani Izziv* **2019**, *30*, 115–128. [CrossRef]
44. De, U.K.; Vupru, V. Location and neighborhood conditions for housing choice and its rental value: Empirical examination in an urban area of North-East India. *Int. J. Hous. Mark. Anal.* **2020**, *10*, 519–538. [CrossRef]
45. Digambar, A.P.; Mazumder, T. Residential location choice: A study of household preferences for the city of Nagpur. *Inst. Town Plan. India J.* **2010**, *7*, 1–19. Available online: http://www.itpi.org.in/uploads/journalfiles/jul1_10.pdf (accessed on 5 November 2021).
46. Rehman, A.; Jamil, F. Impact of urban residential location choice on housing, travel demands and associated costs: Comparative analysis with empirical evidence from Pakistan. *Transp. Res. Interdiscip. Perspect.* **2021**, *10*, 100357. [CrossRef]
47. Subhan, S.; Ahmad, E. The economic and demographic effects on housing tenure choice in Pakistan. *Am. Int. J. Contemp. Res.* **2012**, *2*, 15–24. Available online: http://www.aijcrnet.com/journals/Vol_2_No_7_July_2012/3.pdf (accessed on 2 November 2021).
48. Choudhury, C.F.; Bint Ayaz, S. Why live far?—Insights from modeling residential location choice in Bangladesh. *J. Transp. Geogr.* **2015**, *48*, 1–9. [CrossRef]
49. Shawal, S.; Ferdous, J. Study on the Factors Influencing Residential Location Choice of the Garments Workers of Mirpur, Dhaka City. *Int. J. Chem. Process Eng. Res.* **2014**, *1*, 73–86. [CrossRef]
50. Pakistan Bureau of Statistics. *6th Population and Housing Census—2017*; Pakistan Bureau of Statistics: Islamabad, Pakistan, 2017. Available online: https://www.pbs.gov.pk/content/final-results-census-2017 (accessed on 20 November 2021).
51. Naeem, N.; Ahmad, M. *Residential Location Choice Behaviour: A Case Study of City Hafizabad, Punjab*; Under-Graduation; University of Engineering and Technology: Lahore, Pakistan, 2018.
52. Cochrane, W.G. *Sampling Techniques*, 3rd ed.; Johan Wiley Sons: Hoboken, NJ, USA, 2007; pp. 1–448.

53. Nawaz, M.; Aslam, A.B.; Tariq, F. Evaluation of Government Programs Policies to Eradicate Gender Disparities in Pakistan. *Glob. Reg. Rev.* **2021**, *6*, 22–36. [CrossRef]
54. Prashker, J.; Shiftan, Y.; Hershkovitch-Sarusi, P. Residential choice location, gender and the commute trip to work in Tel Aviv. *J. Transp. Geogr.* **2008**, *16*, 332–341. [CrossRef]
55. Mok, D. Do two-earner households base their choice of residential location on both incomes? *Urban Stud.* **2007**, *44*, 723–750. [CrossRef]
56. Adhikari, B.; Hong, A.; Frank, L.D. Residential relocation, preferences, life events, and travel behavior: A pre-post study. *Res. Transp. Bus. Manag.* **2020**, *36*, 100483. [CrossRef]
57. Klinger, T.; Lanzendorf, M. Moving between mobility cultures: What affects the travel behavior of new residents? *Transportation* **2016**, *43*, 243–271. [CrossRef]
58. Yayboke, E. Five Ways COVID-19 Is Changing Global Migration. Center for Strategic International Studies (CSIS) [Online]. Available online: https://www.csis.org/analysis/five-ways-covid-9-changing-global-migration (accessed on 29 October 2020).
59. Marcucci, E.; Stathopoulos, A.; Rotaris, L.; Danielis, R. Comparing single and joint preferences: A choice experiment on residential location in three-member households. *Environ. Plan. A* **2011**, *43*, 1209–1225. [CrossRef]
60. Beatley, T. Planning for Sustainability in European Cities: A Review of Practice in Leading Cities (2003; updated 2013). In *Sustainable Urban Development Reader*, 3rd ed.; 2014; pp. 440–449. Available online: https://www.taylorfrancis.com/chapters/edit/10.4324/9781315770369-45/planning-sustainability-european-cities-review-practice-leading-cities-2003-updated-2013-timothy-beatley?context=ubx&refId=d94ad759-bf0b-4c23-9981-f9319a8c736a (accessed on 5 December 2021).
61. Abdullah, M.; Ali, N.; Hussain, S.A.; Aslam, A.B.; Javid, M.A. Measuring changes in travel behavior pattern due to COVID-19 in a developing country: A case study of Pakistan. *Transp. Policy* **2021**, *108*, 21–33. [CrossRef]
62. Musselwhite, C.; Avineri, E.; Susilo, Y. Editorial JTH 16–The Coronavirus Disease COVID-19 and implications for transport and health. *J. Transp. Health* **2020**, *16*, 100853. [CrossRef] [PubMed]
63. Takenaka, A.K.; Gaspar, R.; Villafuerte, J.; Narayanan, B. COVID-19 Impact on International Migration, Remittances, and Recipient Households in Developing Asia. 2020. Available online: http://hdl.handle.net/11540/12258 (accessed on 27 November 2021).
64. Scheiner, J.; Holz-Rau, C. Changes in travel mode use after residential relocation: A contribution to mobility biographies. *Transportation* **2013**, *40*, 431–458. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
www.mdpi.com

Sustainability Editorial Office
E-mail: sustainability@mdpi.com
www.mdpi.com/journal/sustainability

MDPI

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-1317-9