Building adaptive smart transport governance using citizen-centric data

Yiqiao Chen

Churchill College

Department of Land Economy

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

July 2022

Declaration

This thesis is submitted according to the requirements of the Degree Committee of Land Economy. It does not exceed the regulation length of 80,000 words including footnotes, references and appendices. It is the result of my own work and includes nothing which is the outcome of work done in collaboration with others, except where specifically indicated in the text and Acknowledgements.

Portions of this work have been published or submitted for review. The following parts of the research are collaborative works with my PhD supervisors:

Chapter Three - **Chen, Y.,** & Silva, E. A. (2021). Smart transport: A comparative analysis using the most used indicators in the literature juxtaposed with interventions in English metropolitan areas. *Transportation research interdisciplinary perspectives*.

Chapter Four - **Chen, Y.,** Silva, E.A. and Reis, J. (Forthcoming). Understanding daily activity-travel sequences of Londoners. Manuscript under review.

The following part of the research is collaborative works with another scholar from our laboratory:

Minor Part of Chapter Five - **Chen, Y**., Niu, H. and Silva, E.A. (2022). The road to recovery: sensing public opinion to reopening measures with social media data in post-lockdown cities. *Cities*.

Yiqiao Chen

Cambridge, 2022

Acknowledgements

I would like to thank all the people who helped me during these four years of PhD study with my greatest gratitude.

First and foremost, I would like to express my deepest gratitude to Prof Elisabete Silva for her enthusiasm, wisdom, and generous support over my PhD years. Without her invaluable advice and continued encouragement, I will not be able to finish this PhD. Moreover, I was given diverse opportunities at the LISA lab, including collaborating with lab members, lecturing students, presenting at conferences, publishing articles and book chapters, coordinating workshops, and the list goes on. I was very fortunate to be in her amazing lab. Additionally, she is more than a supervisor to me. I deeply admire her passion in both academic and personal lives.

I am also extremely grateful to my co-supervisor Dr Jose Reis for his insightful comments and kindly help in accessing data, discussing results, and drafting papers. His patience, intelligence and generosity with his time have contributed to this PhD journey.

I would like to thank two examiners, Dr Stefania Fiorentino and Dr Lun Liu, for taking time and effort to review the thesis and provide valuable comments.

This thesis would not have been possible without the support of many organisations. Many thanks to the Chinese Scholarship Council for the financial support. I would like to acknowledge the UK Data Service, Department for Transport National Travel Survey Team, and the Transport for London for providing the NTS data and the LTDS data, with special thanks to Emily Durrant, Laura Bradshaw, and Hannah Groot for their assistance.

I have benefited immensely from the interactions with PhD colleagues and staff at the Department of Land Economy, particularly at the LISA lab. I would like to thank Dr Haifeng Niu for collaborating with me on social media data analysis. And I would also like to thank Dr Rain Kwon for reviewing this thesis.

Furthermore, I had a great time at Churchill College. The friends, staff and fellows of Churchill College have made Cambridge feel like home.

I would like to thank all my friends in Cambridge and the rest of the world for sharing my happiness in the good times and providing warm encouragement during the hard times.

Special thanks to my boyfriend, Qiuran Peng, for always being there for me. His love, passion, and humour have brought enormous happiness to my life. All the sweet times we spent together have made my life in Cambridge treasured.

Lastly, I would be remiss in not mentioning the dearest (extended) family members: my mother, Chuhong Wang, my uncles, Chuhong Wang and Chupeng Wang, and my cousin, Yunyan Wang. Thank you all for the unconditional love, continuous encouragement, and strong support over all these years. Without the strong support at home, I would not be able to come that far and make it through this PhD.

Summary

With the increasing popularity of the concept “smart city”, many cities have adopted smart governance to address complex socio-economic and spatial issues in urban areas. Smart transport governance is applying innovations in the process of collective decision making in response to the technological and other changes in smart transport development. Governing smart transport, as a key priority in smart cities, faces old and new challenges such as managing complex uncertainties, considering alternative futures, involving citizens and correct analysis of their needs, as well as changing roles of governance. Robust theoretical and practical understandings of smart transport governance are useful for planners and policymakers to address these challenges and transform the urban mobility system towards accessible, sustainable, and innovative futures.

This PhD research explores the complexities in smart transport governance from theoretical, methodological, and practical aspects with a special focus on citizens’ needs. Four gaps in theory, methods, and practice are addressed in six chapters. In Chapter 2, a systematic literature review is performed to enhance the theoretical understanding of smart transport governance and its linkage with complexity theory in cities (CTC) and urban data science (UDS). A citizen-centric adaptive governance framework is proposed. Using the proposed framework to understand specific issues in smart transport governance, Chapters 3-5 conduct empirical studies. Chapter 3 first assesses the existing smart transport governance and development, using a new evaluation framework. Within English metropolitan areas, Greater London ranks first in smart transport development. Chapter 4 zooms into Greater London and applies novel methods to understand citizens’ activity-travel patterns with uncertainties. Typical activity-travel patterns before COVID-19 and the emerging self-organising changes when COVID-19 first hit London are identified. To supply quick insights into

the pandemic's impact on different sub-systems, Chapter 5 senses the public opinion towards different transport sub-systems through real-time social media big data. Dynamic behavioural changes and potential opportunities for smart transport transitions are found.

The outcomes of this research support the idea that CTC and UDS can enhance existing smart transport governance in terms of adaptive planning, robust analysis, and citizen involvement. We have identified and discussed emerging technologies and abrupt crises that add complexity to the urban transport sector on its way to transforming into smart transport. Adaptive understanding with the help of citizen-centric data is crucial for planning uncertain futures. Despite some limitations, the studies can provide theoretical and practical implications for smart transport governance in an increasingly complex world. The study also shows significant potential for future development and further applications of the adaptive governance framework.

Table of Contents

List of Figures

List of Tables

List of Acronyms

API	Application Programming Interface
AVL	automatic vehicle location
ANOVA	Analysis of variance
BAME	Black, Asian and minority ethnic
BEIS	Department for Business, Energy and Industrial Strategy
CA	Combined Authority
CAV	Connected and automated vehicle
CCAV	Centre for Connected and Autonomous Vehicles
CCTV	closed-circuit television
CTC	Complexity Theory in Cities
DfT	Department for Transport
DHD	Dynamic Hamming distance
EV	Electric Vehicles
GL	Greater London
GLA	Greater London Authority
GVA	Gross Value Added
HAM	Hamming distance
ICT	Information and Communications technology
IoT	Internet of Thing
ITS	Intelligent Transport System
LDA	Latent Dirichlet Allocation
LTDS	London Travel Demand Survey
MaaS	Mobility-as-a-Service
NHS	National Health Service
NTS	National Travel Survey
OM	Optimal matching
ONS	Office for National Statistics
PHE	Public Health England
POMS	Profile of Mood States
TfL	Transport for London
UDS	Urban Data Science
WoS	Web of Science

Chapter 1 : Introduction

Cities and other urban areas are where 56.2% of the world's population lives (Habitat, 2020). In the past decades, rapid urbanisation has brought many problems and smart city solutions have been used to solve these problems (Kim, 2022). The rapid development of smart city can be seen in many countries and cities in recent years (Batty, 2013; Kandt and Batty, 2021). With numerous innovations applied in cities, many cities such as London, Singapore, and New York are labelled as "smart cities" (Anthopoulos, 2017).

The concept of "smart city" can be understood from technology-centred, human-centred and hybrid perspectives (Hajek et al., 2022). Smart city approaches have mainly applied information and communications technologies (ICT) for urban growth at the beginning of this century and then used innovations to support sustainable development in the last decade. The goals of smart cities have shifted to sustainability, resilience, and inclusion in recent years (Anthopoulos et al., 2022). A unified smart city definition concerns innovations that are not necessarily ICT-based in urban spaces, with objectives to improve six dimensions (people, governance, economy, mobility, environment and living) (Anthopoulos et al., 2019; Giffinger et al., 2007). This PhD study adopts this unified definition and understands smart cities from a hybrid perspective that considers both technologies and people, focusing on applying innovative methods to solve urban problems, understand citizens' needs, and build adaptive smart cities. This thesis contributes to smart city studies by exploring smart city governance from a holistic perspective. The impacts of technological innovations and citizen science are both discussed.

Smart governance is regarded as the core of the smart city concept and is expected to lead the development of other smart city dimensions (Fernandez-Anez et al., 2018).

In smart cities, smart governance included adopting innovative approaches for decision making process and achieving improved outcomes thorough innovative use of tools (Jiang, 2021). Smart governance contains management and planning in the unified smart city conceptual model (Anthopoulos et al., 2019). Within other key dimensions of smart city, smart transport is one of the priorities in smart city development that requires good governance (Chen and Silva, 2021). Against this background, this PhD thesis particularly focuses on the governance of smart transport in the smart city concept.

1.1 Motivation and objectives

Many cities have adopted smart governance to address complex socio-economic and spatial issues in urban areas. In the smart city era, emerging smart transport innovations can be seen in propulsion, vehicle controls, business models, planning and policies. Transport governance, therefore, faces old and new challenges such as unsustainable mode choices, disturbance of shared bikes, and uncertain impacts of emerging innovations and the COVID-19 pandemic. Smart transport governance that applies innovations in the process of collective decision making and urban mobility management is proposed to address the challenges and uncertainties. Smart transport governance contains a set of strategies, schemes, policies, projects and actions, including integrated ticketing, travel apps, electric vehicles, automated vehicles, and sustainable transport policies (Woods et al., 2017; Harriss and Kearney, 2021). There is a growing body of literature recognise the importance of smart transport governance recently (Docherty et al., 2018; Lyons, 2018; Docherty, 2018; Kumar et al., 2018).

Facing the ever-changing urban mobility system, a primary concern of smart transport governance is the complexity. Understanding complexity has long been a question of great interest in urban and transport studies (Schneider et al., 2021; Liu et al., 2021; Saberi et al., 2017). Many researchers have viewed smart transport as a complex adaptive system and analysed its complex issues, including dynamic contexts of smart mobility, uncertain travel activities and demands, new methodological advancements to unfold complexities, changing roles in governing smart transport, and complex smart transport transitions (Docherty et al., 2018; Pinna et al., 2017a; Farooq et al., 2019; Audouin and Finger, 2018; Pangbourne et al., 2020; Moscholidou and Pangbourne, 2019; Oldbury and Isaksson, 2021). Following this line of thought, this PhD research intends to understand uncertainties in smart transport and support smart transport governance through the complexity perspectives. This thesis particularly pay attention to the interventions and data-driven evidence in smart transport governance.

Most existing studies have only mentioned the complex characteristics of smart transport but have not explicitly incorporated complexity theory (Ribeiro et al., 2021; Docherty et al., 2018; Kester, 2018; Field and Jon, 2021; Moscholidou and Pangbourne, 2019; Icasiano and Taeihagh, 2021). Complexity theory has triggered a "complexity turn" in urban studies, public administration, and political science, with new implications for planning and management (Cairney, 2012; Eppel and Rhodes, 2018; Alexander, 2020; Skrimizea et al., 2019). In the urban domain, previous urban researchers have brought complexity theories into urban studies (Batty, 2010; Portugali, 2012; De Roo and Silva, 2010). We use the term "complexity theory in cities (CTC)" to describe a set of concepts and frameworks from complexity theories to analyse complex urban systems. CTC is an emerging planning theory lately (De Roo and Silva, 2010). So far, however, a systematic understanding of how CTC contributes to smart transport governance is still lacking.

Another key aspect of smart transport governance is the use of data. Data analytics has been widely deployed to generate insights and provide evidence for smart transport management. In knowledge discovery, data science is now known as the fourth scientific trend (Thakuriah et al., 2017; Collins, 2010). Data science has been widely adopted for analysing urban data and supporting data-driven decision making (Kang et al., 2019; Batty, 2019; Bibri, 2018). In this thesis, the term “urban data science (UDS)” is used to refer to a holistic package of data sources, techniques, methods, and knowledge related to the city, with aims to support governance through data-driven decision making (Kang et al., 2019; Batty, 2019; Bibri, 2018). Existing research recognises the critical role of UDS but mainly focuses on the methodological aspect. UDS is more than data analysis using spatial data to understand urban issues (Kang et al., 2019). This indicates a need to further understand the role of UDS in smart transport governance in both theoretical and methodological aspects.

The motivation of this PhD research is to respond to the uncertainties in smart transport governance with an overall hypothesis that CTC and UDS can support smart governance. The research uses English metropolitan areas, particular Greater London, as the empirical cases. The overarching research question is how to enhance the understanding and governance of smart transport through adaptive planning and citizen-centric data analytics. The overarching question can be divided into four main research questions. The four main research questions focus on different aspects of smart transport governance. The research questions and outputs are summarised in Figure 1-1.

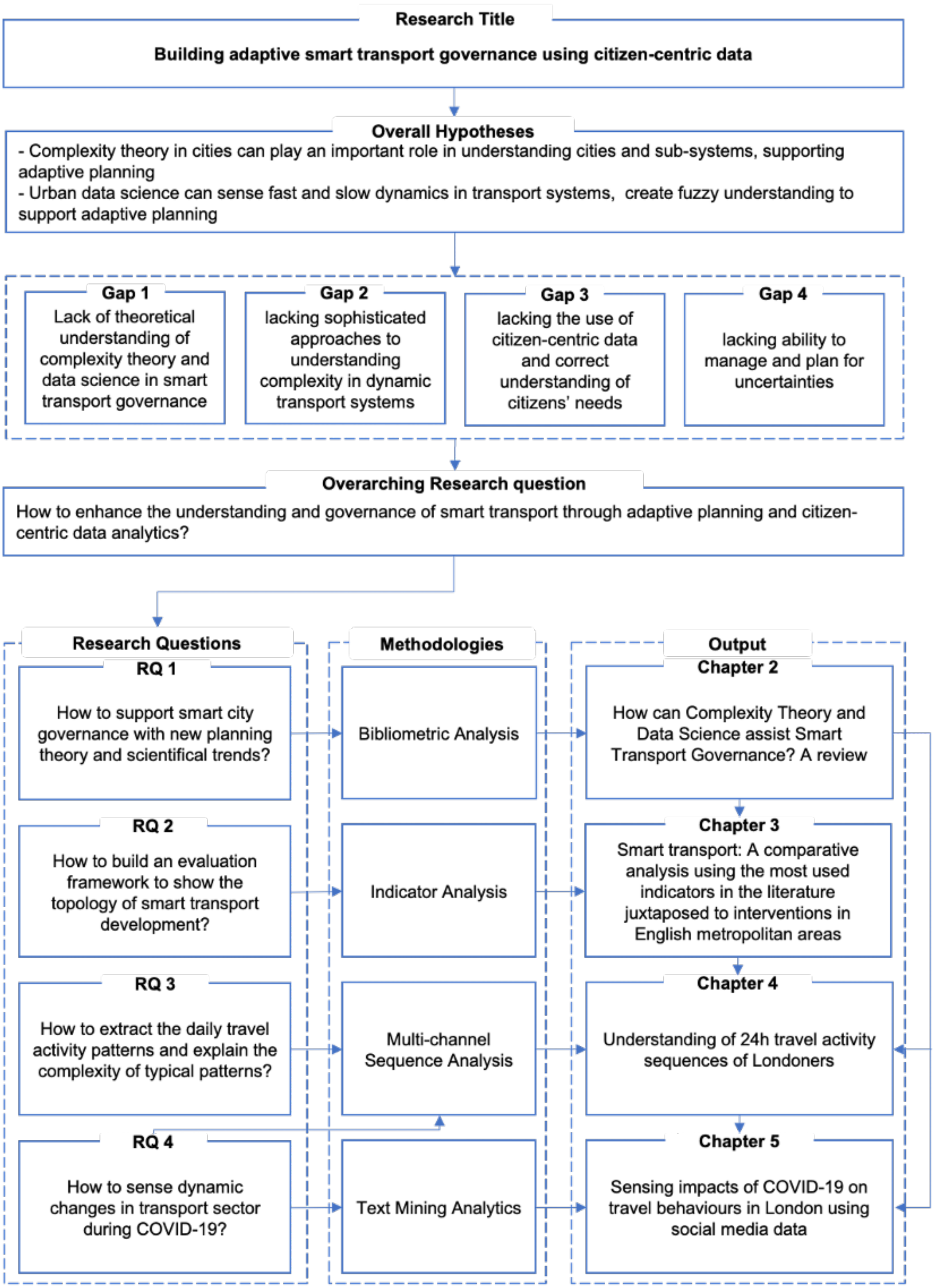

Figure 1-1: Overall research hypotheses, research questions, methodology and outcomes

1.2 Research gaps

Despite a large number of existing studies on smart transport, many research gaps still exist. The gaps are related to multiple facets of research, including theoretical understanding, methodological advancement, and governance practices. This PhD thesis identified four gaps in theories, methodologies, and governance. To fill in these gaps, a set of questions are raised accordingly.

1.2.1 Gaps in theoretical understanding

- Gap 1: lacking theoretical understanding of complexity theory and data science in smart transport governance

An underlying assumption of the smart city is that the innovations are founded on the application of complexity theory and data science (Bibri and Krogstie, 2017b). However, there is not yet a publication that systematically discusses how CTC and UDS assist smart transport governance. Existing studies on smart transport governance have mostly used notions and techniques from the two emerging fields (CTC and UDS) without explicitly discussing the two sets of theories. Many studies have adopted the view of smart transport as a complex adaptive system, analysed complexity characteristics, and pointed to complex transitions with alternative options. However, only a small amount of smart transport literature has directly and explicitly discussed complexity theory in the theoretical part (Ribeiro et al., 2021; Docherty et al., 2018; Kester, 2018; Field and Jon, 2021; Moscholidou and Pangbourne, 2019; Icasiano and Taeihagh, 2021). Although data analytics is widely seen in the methodology part of the reviewed studies on smart transport governance (Sudmant et al., 2021; Nasser et al., 2021), data science is more than analysis and the theoretical aspects are often neglected (Donoho, 2017; Kang et al., 2019; Cleveland, 2014). Rarely have articles systematically explored how UDS can assist smart transport governance in theory. Additionally, the nexus of CTC and UDS are not clear and their linkages with smart transport governance need to be bridged. Thus, we encountered the first research gap when reviewing the recent literature on smart transport governance at an early stage.

1.2.2 Gaps in methodological advancement

When reviewing relevant literature in Chapter 2, two methodological gaps are also found.

- Gap 2: lacking sophisticated approaches to understanding complexity in dynamic transport systems

Despite increasing recognition of the complex scenarios and uncertainties, most existing studies have used different approaches to generate insights from multi-sourced transport data, often with optimisations and high certainties (Anda et al., 2017; Kandt and Batty, 2021). For example, hard clustering algorithms that assign each item to one certain group are widely used in classifying travel patterns while fuzzy clustering methods that consider the non-linear nature and flexibility in data are less seen (Ferraro and Giordani, 2020; Bolin et al., 2014; Li and Lewis, 2016). Methods that consider spatiotemporal dynamics and a level of uncertainties are needed to understand the complexity of smart transport systems (Nasser et al., 2021; Kandt and Batty, 2021). Thus, the first methodological gap concerns the need for robust approaches in investigating complexity, especially in considering uncertainties and providing new insights into both long-lasting and new issues in the dynamic urban transport systems.

- Gap 3: lacking the use of citizen-centric data and correct understanding of citizens' needs

Linked with gap 2, a vast amount of data from multiple sources is generated in smart cities and transport sectors to assist decision-making and service management (Resch and Szell, 2019; Balduini et al., 2019; Kitchin, 2016). However, urban analytics is criticised for being technocratic and uneven (Kitchin, 2019). Data-driven evidence can lead to digital unevenness and misunderstanding of citizens' needs if the information from the transport users and the wider public is neglected (Ma et al., 2018; Mladenovic and Haavisto, 2021; Resch and Szell, 2019). Human-centric data science that understands citizens' needs from data aims for creating value for citizens (Resch and Szell, 2019). The emerging new data to understand citizenry science on the human conditions include: 1) governmental data, 2) data related to official registration/licensing, 3) commercial transactions, 4) internet search records and social media data, 5) tracking data, and 6) image data (Entwisle and Elias, 2013; Milne and Watling, 2019). Thus, there is a gap in including human-centric data science and the correct understanding of citizens' needs in studying smart transport governance (Kang et al., 2019; Resch and Szell, 2019; Sagaris, 2014).

1.2.3 Gaps in governance practice

The last gap is associated with the governance practices and is also identified from the systematic literature review in Chapter 2.

- Gap 4: lacking ability to manage and plan for uncertainties

Facing the ever-emergent, flexible, and spontaneous mobility changes caused by new products, new business models, and the abrupt global pandemic, existing transport governance was criticised for not being smart as it lacks quick and adaptive responses to dynamic changes and cannot make use of opportunities in uncertainties (Kaaristo et al., 2020; Field and Jon, 2021). Additionally, during the complex transitions in urban mobility during the pandemic and in the post-COVID world, many transport authorities showed limitations in managing uncertainties. The need to manage uncertainties is crucial in the current smart transport governance.

During this PhD study, the abrupt outbreak of the COVID-19 pandemic has dramatically changed urban transport systems and created a large extent of uncertainties in the mobility future. Alternative activity routines, modes of travelling, and social norms have emerged. New activity and travel demands in the post-COVID world are unpredictable and the transition/development paths are unknown (Bojovic et al., 2020). Robust understanding is urgently needed to provide evidence for transport management and operation as the impacts of COVID-19 are unknown to all, especially in the early stage (Haken et al., 2021). There is a need to understand the uncertainties brought by COVID-19.

1.3 Research questions and outputs

This thesis conducts theoretical and empirical studies mutually to address the gaps above. Four research questions and twelve sub-questions are pursued. The first research question addresses the theoretical gap, and the last three questions target gaps in methodology and governance. The last three questions are based on the answers to the first research questions. The four research questions are:

1. How to support smart city governance with new planning theory and scientific trends;

2. How to build an evaluation framework to show the topology of smart transport development;

3. How to extract the daily travel activity patterns and explain the complexity of typical patterns; and

4. How to sense the dynamic changes in the transport sector during COVID-19?

To fill in the gap 1, three sub-questions of research question 1 are further set as:

1.1 What are the key themes in smart transport governance studies;

1.2 How can CTC support smart transport governance;

1.3 How can UDS support smart transport governance?

A systematic literature review on relevant literature in the recent five years will be conducted to answer these sub-questions, using bibliometric analysis. The answers to the first research question will be presented in Chapter 2. We will summarise the key themes in existing smart transport governance studies and the implications of CTC and UDS. An integrated framework will be proposed to better understand smart transport issues and support smart transport governance.

Then, to fill in the methodological gaps 2, we set the second question as “how to build an evaluation framework to show the topology of smart transport development” to provide a big picture of smart transport development in the UK, with three sub-questions:

2.1 How do English metropolises govern smart transport in terms of interventions;

2.2 What are the most common and important indicators and indices to examine smart transport;

2.3 What are the smart transport developments in the English metropolises?

A systematic literature review of existing indicators will be conducted, followed by a search for new indicators to build a new assessment framework. Multi-sourced data will be collected and analysed in the indicator analysis. Seven synthesised indices will be calculated to show the smart transport development in different cases. Chapter 3 will answer the second question and sub-questions.

Building on the empirical findings of the second question, the third research question mainly addresses the methodological gaps 2 and 3, using Greater London as the case area. Three sub-questions are designed:

3.1 How to identify representative patterns of daily activity-travel sequences;

3.2 What explains the complexity of Londoners' daily activity-travel patterns based on their travel sequences and socio-demographic profiles;

3.3 What is the difference in Londoners' daily activity-travel sequences before and during the early stage of the pandemic?

We will unfold the complexity of London's urban transport system through an adaptive understanding of the daily activity patterns of citizens, using fuzzy multi-channel sequence analysis methods. The answers to the third research question will be in Chapter 4.

During this PhD study, the abrupt global health crisis of COVID-19 has extensively changed the urban transport system and brought uncertainties in cities. Under this circumstance, we raise the fourth question to specifically address governance gap 4. Three sub-questions are set as:

4.1 What is the public sentiment towards the pandemic itself and the transport sectors during COVID-19;

4.2 What are the impacts of COVID-19 on different transport sub-systems;

4.3 What are the opportunities in the uncertainties?

To answer the questions above, we will deploy real-time social media big data to illustrate the fast changes under governance, underlying reasons for changes in mobility patterns, and potential long-term effects. The answers will mainly be in Chapter 5.

1.4 Summary of methodological approaches

This thesis uses mixed research methods to answer four research questions above, as described below.

- Systematic Literature review

This research first systematically reviews publications relevant to smart transport governance, complexities theory in cities, and UDS to map the key domains of smart transport governance studies and to link them with complexity theory and data science. We also systematically review relevant articles on smart transport indicators to identify the most used indicators. Bibliometric analysis is conducted to analyse the metadata of reviewed publications. In bibliometric analysis, maps of keywords' co-occurrence, bibliographic coupling and co-citation are generated to show the main themes in relevant studies, interrelationships among articles, and influential co-cited articles based on bibliographic data.

- Case studies

The empirical studies select ten Combined Authorities in England and Greater London as case studies, using multi-sourced data from these metropolitan areas. The existing ten CAs are Cambridgeshire and Peterborough, Greater Manchester, Liverpool City Region, North of Tyne, Sheffield City Region (renamed as South Yorkshire in 2021), Tees Valley, West Midlands, West of England, West Yorkshire, and North East Combined Authorities. Greater London is often seen as the first CA with devolution power and used as a beacon in analysing CAs.

- Indicator analysis

An indicator or index can represent a specifically evaluable phenomenon through suitable measurements (Lopez-Carreiro and Monzon, 2018; Kitchin et al., 2015; Battarra et al., 2018a). To systematically assess the development of smart transport in the case studies above, indicator analysis is conducted to identify trends and emerging dynamics in smart transport. The disaggregate indicators are further combined into indices to show the overall performance of important elements in smart transport systems.

- Spatiotemporal sequence analysis

To find space-time patterns of citizens' activities and trips, we apply multi-channel sequence analysis. The method can represent an individual's movement as sequences in multiple dimensions and each sequence in a dimension shows a consistent state (e.g., location state, activity state, trip state) of the movement through time (Brum-Bastos et al., 2018). It is a powerful analytical framework to extract human activity and mobility patterns (Hafezi et al., 2017; Zhang et al., 2019). Similar spatiotemporal patterns can be identified through comparing, aligning, clustering, classifying, and profiling activity-travel sequences.

- Text mining methods

Social media big data from Twitter is used to supply the quasi-real-time understanding of the COVID-19 impacts on urban transport. The unstructured textual data is first cleaned. We then use the AFINN sentiment analysis and the Profile of Mood States emotion detection model to identify the sentiment polarity and emotion states of tweets. Dynamic Latent Dirichlet Allocation topic models that can group themes with possibilities and consider temporal changes are then built to explore semantic structure from texts and how topics changed over time.

1.5 Contributions and Innovations of research

This PhD research provides relevant contributions to learning by:

- producing a comprehensive portfolio of main themes in smart transport governance literature and supplying the existing theoretical understanding with CTC and UDS (Chapter 2),

- proposing a new citizen-centric adaptive governance framework for smart transport governance in the uncertain world (Chapters 2),

- building a robust and up-to-date evaluation framework to compare smart transport development in English metropolitan areas (Chapter 3),

- providing broad insights into the smart transport governance in the English metropolitan areas (Chapter 3),
- employing innovative data mining approaches to extract adaptive understandings of mobility patterns and behavioural changes (Chapters 4 and 5),
- identifying the typical sequential travel patterns and socio-demographic determinants in Greater London (Chapter 4),
- sensing the COVID-19 impacts on transport sectors through social media mining (Chapter 5),
- discussing the potential of linking big and "small data" in assisting smart transport governance (Chapter 5).

1.6 Thesis structure

This PhD research is divided into six chapters, composed of theoretical and empirical parts. After this introduction, Chapter 2 provides a systematic literature review of smart transport governance, CTC, and UDS.

Based on the theoretical understanding, the empirical part of this thesis contains three empirical studies in three chapters. Chapter 3 conducts the first empirical study that focuses on benchmarking smart transport development and governance, using indicator analysis. To further understand the governance and complexity in the smartest city in the transport domain - Greater London, Chapter 4 analyses the daily activity-travel patterns of Londoners through innovative fuzzy multi-channel sequence analysis. To rapidly sense the impacts of COVID-19 on Londoners' travel behaviours and urban transport in London, Chapter 5 further deploys real-time social media big data to sense the public opinion on different transport systems.

Based on the theoretical and empirical findings, Chapter 6 finally summarises the main finding of this thesis, emphasising the adaptive governance framework for smart transport, key governance implications for empirical studies, and pointing out the directions for future research.

Chapter 2 : How can Complexity Theory and Data Science assist Smart Transport Governance? A Review

2.1 Introduction

As mentioned in Chapter 1, smart governance is a key element in the smart city concept and it can lead other subsystems (Fernandez-Anez et al., 2018; Meijer and Bolivar, 2016). Another key dimension in smart city concept is smart transport. Smart transport, as a priority in smart city development, needs good governance to achieve the overall smart city goals of inclusion, sustainability, and better quality of life (Docherty et al., 2018).

Both smart cities and the transport sector can be seen as complex systems with constantly changes, fast or slow. Increasing number of research explores the integration of complexity properties in understanding and simulating the complex transport systems with the help of conceptual and methodological frameworks from complexity theories (Avineri, 2016). To better govern these complex systems, we need to understand them from a complexity perspective. Complexity theories, particularly the CTC, have the potential to provide theoretical understanding of the inherent dynamics of smart transport governance in smart cities (Ekman, 2018; Colding et al., 2020; Bibri, 2018). Although complexity is widely mentioned in the smart transport governance literature, the question of how complexity theory can support smart transport governance remains unclear.

Meanwhile, it is widely acknowledged that the emergence of data science has changed the paradigm of knowledge discovery (Collins, 2010). We are now in a time of abundant data and techniques, allowing researchers and practitioners to advance their understanding of old and new phenomena in cities, particularly the data-rich and even "data deluge" smart cities (Singleton and Arribas-Bel, 2021; Batty, 2013). UDS has enriched our understanding of how urban systems function, especially in the short-term (Batty, 2013). Urban transport has been one of the focuses in academia and industries during the rise of big data. UDS can provide new understandings of smart transport governance (Batty, 2013; Anda et al., 2017). Although the importance of data is widely acknowledged, existing studies have not dealt with the implications of data science on smart transport governance in recent years when innovations rapidly emerge. Thus, it is necessary to explore the implications of data science on the increasingly complex smart transport governance.

To understand the recent literature landscape of smart transport governance, and how CTC and UDS can further support smart transport governance, we conduct literature searches on the Web of Science to include the most relevant and meaningful articles from the previous five years (2017-2021). As these fields are changing fast (particularly influenced by quick technology adoptions), we choose the papers from the recent five years to show the new academic focuses and recent literature landscapes. Through analysing the reference lists of the reviewed papers, the seminal papers from other years are identified and utilised in explaining ideas. The literature review adopts systematic methods with a criterion-based selection process (Wolfswinkel et al., 2013; Ruhlandt, 2018). The seven procedures of literature search are: 1) choose databases, 2) select keywords, 3) specify filter type, 4) remove duplications, 5) filter by keywords, 6) refine by full text, and 7) add forward/ backward citations (Ruhlandt, 2018). For smart transport governance, 58 peer-reviewed articles are left for review. We analyse 110 articles on CTC and 255 on UDS. The bibliometric analysis to explore relationships among publications is used for extracting the conceptual and intellectual structure of reviewed literature (Aria and Cuccurullo, 2017). The metadata of these publications, including citations, abstracts, keywords, and other information, is measured (Van Eck and Waltman, 2009). Through mapping keyword co-occurrences, we can reveal the recent trends in the previous five years. The co-citation maps that analyse cited references in reviewed papers can show the classic and important articles in the field from all years. We also use these highly co-cited papers from other periods to reinforce our arguments in this chapter.

The chapter is organised as follows. Section 2.2 reviews papers on smart transport governance from 2017 to 2021, using bibliometric analysis. Section 2.3 introduces CTC and its guidance on smart city governance. Section 2.4 presents UDS and the main implications of governing smart cities. Section 2.5 discusses the nexus between CTC and UDS in smart transport governance, points to the citizen-centric adaptive governance framework, suggests a multidisciplinary approach to support smart governance, and discusses further directions. Section 6 concludes the paper by revisiting key findings.

2.2 Smart transport governance

In smart city concept, smart city governance is applying innovations in the process by which governments and stakeholders collectively decide how to plan and manage urban areas. As summarised in Figure 2-1, a smart governance framework contains key components of smart city governance, measurements, global trends, contextual factors, and the outcomes (Bolivar and Meijer, 2016; Ruhlandt, 2018; Fernandez-Anez et al., 2018).

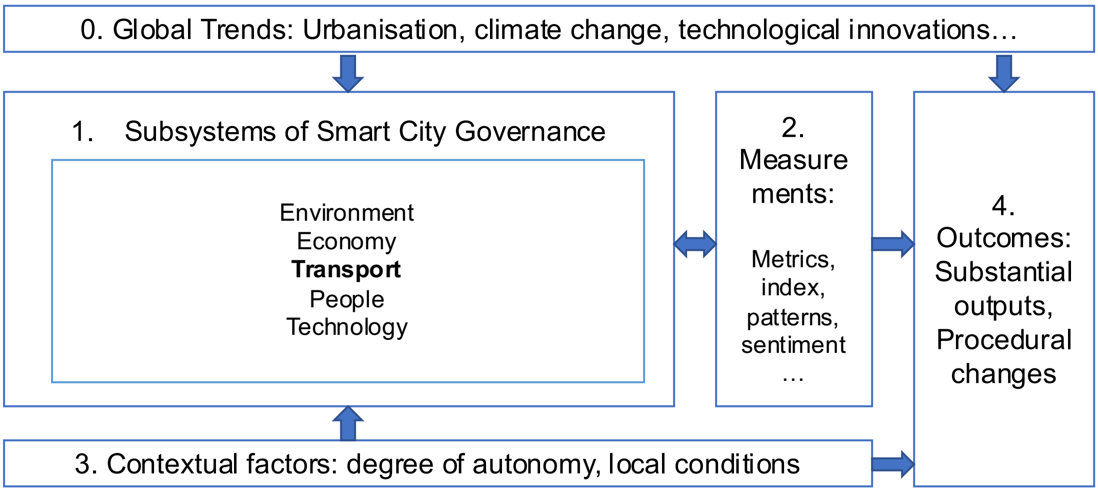

Figure 2-1: Smart city governance framework

Smart transport is a crucial focus in smart city governance as the transport system adopts technological advances faster than other systems (Fonzone et al., 2018). Smart transport is one of the priorities in smart city strategies. For example, transport is a key area in ongoing smart strategies in Greater London, Bristol (in West of England), Greater Manchester, and Birmingham (in West Midlands) in the UK (Chen and Silva, 2021; Woods et al., 2017). In this study, smart transport refers to using innovations to improve connectivity in towns and cities towards affordable, effective, attractive and sustainable mobility futures (Lyons, 2018).

Although smart transport is believed to fulfil these objectives, it needs to be carefully governed to ensure and enhance sustainability because advanced technologies in the transport sector may not bring added public value, as, for instance, the automobility transition in the 20th century has shown that poorly transport governance can result in huge adverse impacts (e.g., air pollution) (Docherty et al., 2018). Without careful governance of smart transport in soft institutions, hard infrastructures with technological advances, and sufficient involvement of users, the public values of smart transport products and interventions may not result in what people expected. For example, researchers have found that the reduction in carbon emission brought by electric vehicles (EV) is questionable because EV reinforce the current automobile paradigm and can reduce cycling rates and the use of public transport, which eventually will increase the carbon emission in cities (Kester, 2018; Friis, 2020). Thus, the good governance of smart transport in the smart city concept is worth investigating.

We conducted a bibliometric analysis for smart transport governance, using the Web of Science (WoS) Core Collection as the main database. The keywords of “smart transport governance” and “smart mobility governance” were used to search for relevant literature. We extracted 247 articles from WoS. After a seven-step selection, we analysed 58 articles published by 178 authors, with a 30% annual growth rate. The most dominant authors are Lyons G, Docherty I, Garau C, Pangbourne K etc. *Sustainability, Transportation Research Part A – Policy and Practice, Transport Policy,* and *Cities* are the main sources. Interdisciplinary knowledge and approaches have been applied in the literature. The main fields of study consist of transport, urban studies, public administration, environmental science and ecology, engineering, computer science, geography, and developing studies.

The conceptual structure of recent literature can be categorised into four main clusters, as shown in Figure 2-2. The red cluster concerns complex transitions, uncertainties, and options in managing future smart mobility such as Mobility-as-Service (MaaS). The green cluster is about methodologies in existing smart transport analysis such as big data analytics. The blue cluster discusses the main topics, including smart transport characteristics, travel activities and demands, and different roles of governance. The yellow group links smart mobility with other subsystems such as environment and economy as well as the broader urban planning.

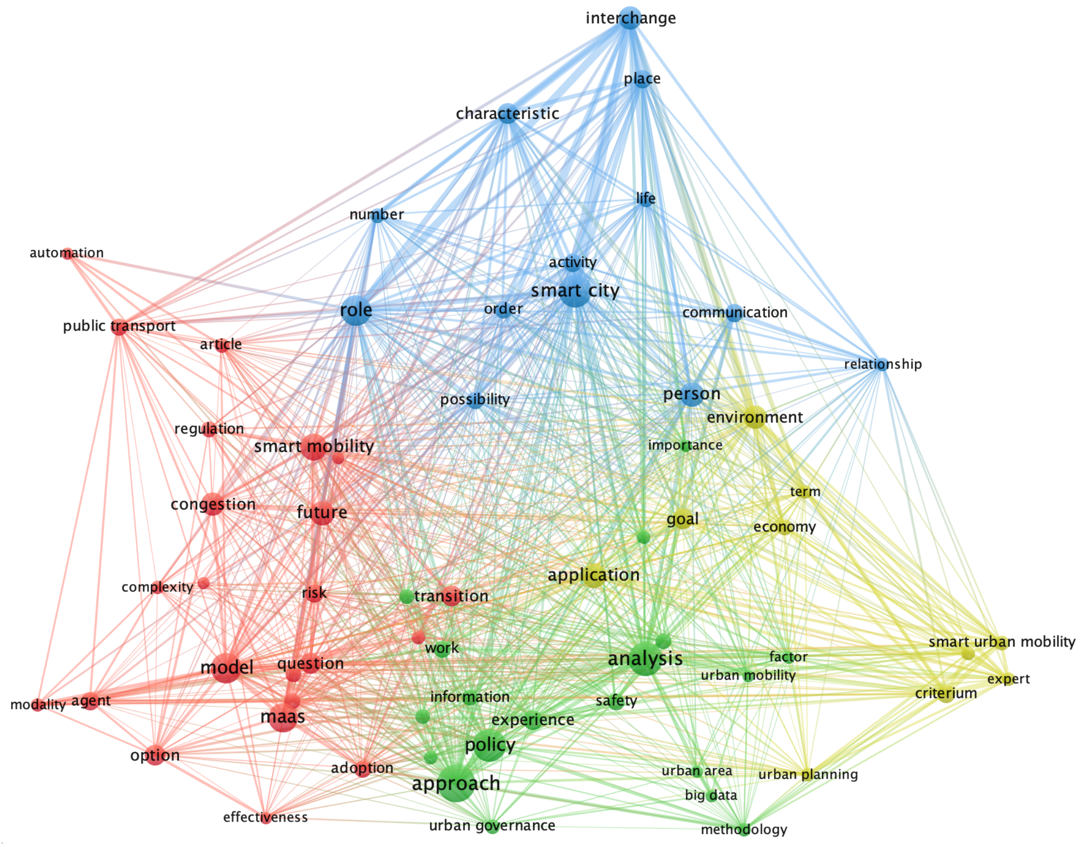

Figure 2-2: Co-occurrence network based on texts from titles and abstracts

The red cluster is associated with the theoretical understanding of smart transport governance. The main theories in the literature are from: 1) technology and innovation studies, 2) political science and public policy views, and 3) complexity theories. Many authors have used socio-technical transition theory from the first theoretical category (Mladenovic and Haavisto, 2021; Pangbourne et al., 2020; Docherty et al., 2018; Girones et al., 2019). The multi-level perspective of socio-technical transition theory is the most used analytical framework in smart transport literature. The analytical framework contains macro-level contextual environments (landscapes), meso-level dominant structure and practices (regimes), as well as micro-level innovations (niches) (Geels, 2012). Within the technology and innovation theories, other theories such as diffusion of innovations, collaborative and institutional innovations theories have also been used. They are mainly included as complements to socio-technical transition theory (Mladenovic and Haavisto, 2021).

A smaller number of researchers have investigated the smart mobility issues from political science and public policy perspectives, using theories such as contemporary political theory, institution theory, collaborative governance theory, and regulation theory (Field and Jon, 2021; Mladenovic and Haavisto, 2021; Ma et al., 2018; Friis, 2020).

Lastly, complexity theories can be seen explicitly or implicitly in many studies. The mobility system is widely recognised as a non-linear complex system and complexity notions and planning rationales are commonly used in discussions (Ribeiro et al., 2021; Docherty et al., 2018; Kester, 2018; Field and Jon, 2021; Moscholidou and Pangbourne, 2019; Icasiano and Taeihagh, 2021). The complexity lens is increasingly integrated in examining and governing complex mobility systems, including dynamic macro/micro contexts, adaptive capacity, and non-linear transition management (Kester, 2018; Friis, 2020; Ribeiro et al., 2021).

The final theoretical frameworks used in reviewed articles are often a mix of different theories. For example, socio-technical transition theory has been refined by political studies through adding roles of governance of new public management and by complexity theories through seeing smart transport as a complex system (Docherty et al., 2018). In the mixed framework, adaptive capacity in responding to changing circumstances has been regarded as an element in successful non-linear transition management (Smith et al., 2005; Girones et al., 2019).

The green cluster is linked with multidisciplinary methodologies for analysing issues in smart transport. Authors have deployed various data such as interviews, surveys, government reports, Origin-destination data, and distance matrix data (Billones et al., 2021; Vrscaj et al., 2021; Faber et al., 2018; Guo et al., 2020). Main methodologies include cluster analysis, indicator analysis, text mining, machine learning and living lab experiments (Sjoman et al., 2020; Cellina et al., 2020; Garau et al., 2019; Yatskiv et al., 2018; Mounce et al., 2020). Among all methods, big data analytics have been suggested to provide real-time detailed information and advanced sophisticated techniques for mining “small data” (e.g., surveys) have been recommended to extract more meaningful insights (Sudmant et al., 2021; Nasser et al., 2021). However, data-driven approaches often come with concerns such as data privatisation, digital exclusion, and security issues (Hussain et al., 2020; Whitelaw et al., 2020). Researchers have called for open data, robust and advanced analytics, as well as correct analysis of citizens’ needs (Stilgoe, 2018; Abbas et al., 2021; Nasser et al., 2021; Gonzalez et al., 2020). We found methodological gaps of lacking robust approaches to understand the dynamics in smart transport development as well as lacking citizen-centric data mining and correct understanding of citizens’ needs.

The blue cluster concerns the main tasks in smart transport governance. In the recent studies, understanding the dynamic roles of governance or changing roles of transport authorities, assessing smart transport products and interventions, modelling travel demands, improving collaborative governance, exploring dynamic wider and local contexts, and inventing uncertain future mobility through adaptive responses are important tasks (Docherty et al., 2018; Pinna et al., 2017a; Farooq et al., 2019; Audouin and Finger, 2018; Pangbourne et al., 2020; Moscholidou and Pangbourne, 2019; Oldbury and Isaksson, 2021). Within the governance issues, gaps of lacking abilities to manage the uncertain future have been identified. The thesis will address the governance gap and analyse some of the important tasks in the empirical chapters.

The yellow group includes wider interrelated contexts. Good governance in smart transport can result in better smart environment and economy in the wider smart city

(Docherty et al., 2018; Aleta et al., 2017). An undesirable transport system in smart city will also harm the other smart domains such as living and people. The governance in all other interrelated systems can also influence the smart transport system (Abu-Rayash and Dincer, 2021).

2.3 CTC for smart transport governance

As highlighted in section 2.2, smart transport governance has been increasingly analysed through complexity theory. A small number of smart transport literature have directly and explicitly used complexity theory when analysing complex mobility systems while a much wider range of studies have only highlighted the complex system characteristics of smart transport without mentioning complexity theory. Thus, we turned to a broader literature on complexity theory to draw insights to support smart transport governance, including articles from the previous five years and earlier seminal papers that are highly co-cited in the reviewed literature.

Breaking from reductive approaches (that break urban systems into individual components with few weak and linear interactions) for sector by sector in theories and practices, Complexity theory studies the interactions among individual components in a system and the changes of both components and the whole system over time (Lawrenze et al, 2018). It is an umbrella term for a set of concepts and frameworks to analyse complex systems such as cities and transport systems (Alexander, 2020; Cairney, 2012). Complexity is a state between randomness and order (Alexander, 2020; Byrne and Callaghan, 2013). A complex system composes many interrelated components and interactions, with the key features of non-linearity, self-organisation, emergence, path dependence, adaptation, and uncertainty (Wallentin, 2020). Cities and urban transport systems, as complex systems, have many heterogeneous agents and components as well as non-linear interactions within the systems (Alberti et al, 2018).

The first linkage between complexity and cities is the description by Jacobs (1961), stating that cities are “problems of organised complexity”. Physicists such as Prigogine (1978) and Allen (1997) applied notions in complexity theory to describe urban systems. More urban researchers then analysed various aspects of the relationships among complexity, cities, planning and management. Key authors of the highly citied literature included Michael Batty, Juval Portugali, Gert de Roo and Ward Rauws. Michael Batty is a pioneer of thinking cities as complex systems. Batty (2010) published the book *New Science of Cities*, suggesting cities can be understood as networks and flows. He introduced bottom-up evolutionary models that simulate non-linear interactions and flows in cities and pointed out future planning as collective actions. He has introduced non-linearity and complexity into spatial analysis and planning (De Roo et al., 2020). Innes and Booher (2010) proposed collaborative planning approaches to manage complexity and argued that complexity can be a new direction of collaborative planning. De Roo and Silva (2010) published the book *A Planner's Encounter with Complexity*. The book presented planners’ views on complexity and the complexity of theory of planning. De Roo argued that planning theory is in a crisis as the previous two planning theory paradigms are either from a technical rationale (i.e., system planning) or from a communicate rationale (i.e., collaborative planning). However, most planning issues are between the two opposing rationales of planning above, involving a mix of certainty and uncertainty, and required a complexity perspective to understand the ‘in-between’ fuzzy environment (De Roo and Rauws, 2012). Portugali et al. (2012) coined the term “Complexity Theory of Cities” (CTC) in his book *Complexity Theories of Cities Have Come of Age*. Beyond using complexity theories to solve urban problems, CTC focuses on urban dynamics (Portugali et al., 2012). Portugali (2020) particularly focuses on self-organisation within cities and highlights complex adaptive behaviours in cities. CTC has a long history of exploring how cities work but its implications for governance have been less discussed (Portugali et al., 2012: 4).

CTC has reshaped the understanding of cities, transport systems and governance in several aspects. Firstly, the dynamic characteristics of the urban systems are recognised. Non-linearity is the most important feature in complex urban systems (Bibri, 2018). A minor change can lead to unexpected changes and transitions in systems while a huge effort may barely have an impact on systems (Rauws, 2017). Inspired by other complex systems (e.g., biological systems), the dynamic properties of emergence, self-organisation, adaptation, path-dependency, transition, and co-evolution are brought into the understanding of complex urban systems and the transport domain (Bibri, 2018; Sengupta et al., 2016). Sengupta et al. (2016) argued that temporal dynamics is a crucial dimension in linking urban management with the wider complexity notions above. Apart from temporal dynamics, spatial dynamics is another essential aspect as urban and transport issues should be analysed on the right spatial scale (Perveen et al., 2020). Thus, to plan and govern the ever-changing urban transport systems, we need to investigate the dynamic properties temporally and spatially. Through understanding of these dynamic concepts, urban/spatial planners and practitioners can influence the potential paths and evolutions of urban systems towards better directions (Geyer and Cairney, 2015).

Secondly, we rethink the equilibrium and predictability of urban systems. Cities and the subsystems such as transport are not built in a benign and close environment so they may not return to equilibrium. As random insignificant factors may have unexpected and non-linear effects on systems, the complex systems are unpredictable. Under this circumstance, we can only provide “good enough” predictions rather than perfect predictions when analysing complex situations

(Alexander, 2020; Haken et al., 2021).

Thirdly, cities and transport systems can be seen as complex adaptive systems (CAS) (Batty 2012, Skrimizea *et al.* 2019). The CAS is an entity with internal robustness and in a state between order and chaos. The complex adaptive urban/transport systems are open systems that exchange energy and information within the systems and with wider environments, which are more like organisms (De Roo, 2020; Rauws, 2017). Developments in urban transport systems are dynamic processes of "becoming" as the consistent volatilities in cities exist (De Roo, 2016). The nature of urban systems and their underlying processes are continuously changing, rapidly or slowly (Rauws, 2017; Folke, 2006). The developing/transition paths of CAS are sensitive to initial conditions and changes (De Roo, 2020; Rauws, 2017). Therefore, complex adaptive urban transport systems are constantly changing, reorganising, and adapting to better fit the dynamic environments. We can make use of the conditions to partly change the system in the planning in practice.

Lastly, adaptive governance approaches have the potential to manage uncertainties in urban and transport developments. Uncertainties exist in urban systems and unforeseen development trajectories can occur in complex systems and governance processes (Sengupta et al., 2016; Eppel and Rhodes, 2018). The uncertain futures are latent possibilities of collective co-evolution. During uncertain developments, urban actors can grasp emergent opportunities that occur during uncertainties through learning by doing, multi-actor collaborations, and creating/influencing conditions for desirable paths (Sengupta et al., 2016; Eppel and Rhodes, 2018; Rauws, 2017; Kato and Ahern, 2008). Among these actors, decision makers and people working in the public sector need adaptive mindsets to respond to changing circumstances, expectedly and unexpectedly. Planners are transition managers and trend watchers rather than technical experts (De Roo, 2012; Geyer and Cairney, 2015). Embracing the uncertainties, Adaptive governance can enhance responding capacity in managing complexity (De Roo, 2016; Batty, 2012; De Roo and Silva, 2010; Sengupta et al., 2016; Eppel and Rhodes, 2018). The adaptive approach can influence the conditions of urban areas without defining a certain future configuration or actor relationships, allowing unplanned self-organisation and spontaneous co-evolution to emerge with foreseen and unforeseen changes at different spatial and temporal levels (Rauws, 2017). The adaptive planning rationale in governing the uncertain future transport in smart cities should be highlighted.

When it comes to smart transport governance, smart technologies are making cities and transport systems more complex (Batty et al., 2012). Building on the notions of CTC, smart governance needs new forms of policy analysis and planning in the digital era (Batty et al., 2012). As we cannot generate a perfect prediction, analysts should turn to robust analysis and pluralistic styles of models and simulations for understanding complex issues adaptively (Batty and Marshall, 2012; Haken et al., 2021). The rationale of adaptive planning that accepts uncertainties and multiple future directions based on an adaptive understanding of the complex system should be highlighted. To provide adaptive understanding, methods from complexity sciences such as network analysis, fuzzy logic, cellular automata, and agent-based models can be used to unfold the complexity on the right spatiotemporal scale (Portugali et al., 2012: 3). Additionally, complex knowledge from the citizenry and multi-actor collaborations in decision making should be ensured in smart governance (Sagaris, 2014; Ekman, 2018).

To sum up, the key implications from CTC to support smart transport governance are: 1) accepting latent possibilities of uncertain futures (Sengupta et al., 2016; Eppel and Rhodes, 2018), 2) exploiting opportunities in uncertainties (Rauws, 2017), 3) understanding complex issues with interdisciplinary knowledge and approaches (Batty and Marshall, 2012; Haken et al., 2021), 4) using holistic/robust analysis to support decision making (Batty and Marshall, 2012; Haken et al., 2021), 5) increasing the responsiveness to changes through adaptive governance (De Roo, 2016; Batty, 2012; De Roo and Silva, 2010; Sengupta et al., 2016; Eppel and Rhodes, 2018), 6) enhancing adaptivity in new institutional frameworks (Geyer and Cairney, 2015), and 7) adjusting governance to temporal and spatial dynamics (Perveen et al., 2020).

2.4 UDS for smart transport governance

As shown in the green cluster in section 2.2, multi-sourced data and analytics have been widely discussed in smart transport literature. Additionally, much research within CTC has also proposed new data analysis methods to understand the dynamic properties, patterns, and behaviours of complex systems (Bibri, 2018). The emerging data science can enhance both existing smart transport studies and its integration with CTC. To draw insights for increasingly complex smart transport governance, a broader literature on data science was reviewed. These studies were mainly from 2017 to 2021 and more classic articles were included to reinforce our ideas. The classic articles from other years were identified in the reference lists of reviewed studies from the recent five years.

The emerging field of data science is regarded as the fourth paradigm of science, after empirical, theoretical and computational science (Bibri, 2019; Hey et al., 2009). It is the study of extracting knowledge from data through multidisciplinary principles, processes, and techniques in an evidence-based manner (Dhar, 2013; Donoho, 2017; Bibri, 2018). Data science can enhance the accuracy and validity of science through scientific data analysis, but it is beyond data analysis (Donoho, 2017). The six key domains in data science are: 1) data collection, wrangling, and exploration, 2) data representation and transformation, 3) computing with data in programming languages, 4) data modelling, 5) data visualisation and interpretation, and 6) science and theory about data science (Donoho, 2017; Kang et al., 2019; Cleveland, 2014).

In the urban domain, UDS refers to using data science to understand cities, with the aims to support governance through data-driven decision making (Kang et al., 2019; Batty, 2019; Bibri, 2018). UDS is a holistic package of data sources, techniques, methods, and knowledge related to the city (Kang et al., 2019). To work on UDS, both knowledge from data and urban sciences are required, including statistics, machine learning, and knowledge of urban economy, planning, transport, and environment (Kontokosta, 2021; Katakis, 2015).

UDS has provided new data, techniques, and principles to understand cities, transport systems, and urban transport management and outcomes. Firstly, new data has generated more detailed, sophisticated, large-scale, fine-grained, and real-time findings on urban dynamics and human activities (Bibri, 2019; Anda et al., 2017; Bettencourt, 2014). In UDS, multi-sourced data has been deployed, including data from mobile phone companies, transport service providers, mobile phone operators, social media platforms, and government authorities (Resch and Szell, 2019; Balduini et al., 2019; Kitchin, 2016). Emerging big data can overcome the limitations of conventional “small data” (e.g., Census, surveys, interviews) such as smaller samples and long updated time (Anda et al., 2017; Bibri, 2019). Planners and practitioners can deploy new datasets to support decision making.

Secondly, there is an increasing need to link big data and “small data” – use big data analytics to supply “small data” results or vice versa (Anda et al., 2017; Kandt and Batty, 2021; Hong et al., 2022). It should be noted that “small data” refers to more conventional data and it may not actually be small in its 5Vs (i.e., velocity, volume, value, variety, or veracity) (Ishwarappa and Anuradha, 2014; Hong et al., 2022). For instance, the volume of Census microdata can be larger than some big data samples. The importance of “small data” is also highlighted in UDS (Kang et al., 2019; Silva et al., 2021; Hong et al., 2022; Kandt and Batty, 2021). Linking big data and “small data” can provide comprehensive and robust evidence for decision making.

Thirdly, the human-centric data and citizenry science is highlighted in the knowledge discovery processes. Researchers have called for shared and open data to generate the public value of data and emphasised the importance of data from citizens (Kang et al., 2019; Resch and Szell, 2019). Mining the citizen-centric data sources with suitable analytical tools will allow researchers to find the correct citizens’ needs. The emerging new data to support citizenry science include: 1) governmental data, 2) data related to official registration/licensing, 3) commercial transactions, 4) internet search records and social media data, 5) tracking data, and 6) image data (Entwisle and Elias, 2013; Milne and Watling, 2019). Additionally, new concerns on ethical issues, including datafication and privacy, data use, sharing and repurposing are raised in the data-rich environment (Kitchin, 2016; Dennis et al., 2019). Kitchin (2016) suggested that the ethical aspects should be investigated, and governments should take a proactive role in protecting citizens.

Fourthly, we rethink the spatiotemporal dynamics in cities as UDS can provide insights from long-term to real-time dynamics in multiple spatial scales. Smart cities and transport are more or less being shaped and managed through UDS. UDS can help researchers find new characteristics and patterns of fast urban dynamics (Kandt and Batty, 2021). One of the significant contributions of big data is that we can learn about changing patterns and processes from yearly to minutes by minutes and even seconds by seconds (Gong et al., 2020; Silva et al., 2021; Batty, 2013). Based on reliable patterns and interpretation, results from quasi-real-time data can inform policymaking

(Kandt and Batty, 2021). For example, facing the emergent health crisis of COVID-19, UDS has provided many quick insights in the early stage (Wissel et al., 2020; Xu et al., 2021). However, Kitchin (2019) criticised that urban analytics over relies on real-time big data in the short-term and the governance based on real-time approaches can be technocratic and uneven. Urban data scientists need to have a deeper understanding of the role of real-time and space-time relationships in cities and transport systems (Kitchin, 2019). Urban and transport dynamics should be investigated on a suitable spatiotemporal scale.

Fifthly, new techniques and principles are applied in analysing urban transport issues. UDS extracts insightful knowledge from urban datasets, uncovers important attributes of urban actors and entities, and evaluates the results of smart solutions (Bibri, 2018). Old tasks in cities can now be understood through new data perspectives. Advanced urban analytics and data-driven simulations are used to identify urban land-use changes on finer scales, mitigate air pollution and energy consumption with detailed insights, analyse human activity patterns and manage urban traffic in real-time etc (Gao and O'Neill, 2020; Landrigan et al., 2018; Benavente-Peces and Ibadah, 2020; Pasichnyi et al., 2019; Szczepanek, 2020; Gong et al., 2020). UDS can formulate data solutions to solve issues in complex systems.

Lastly, we rethink the role of data and theory in knowledge discovery. Many studies we reviewed in section 2.2 applied grounded theory, which develops theory from data rather than applying theory to data (Nikolaeva et al., 2019). The increasing usage of data science has been criticised for neglecting theory and overemphasising the data mining process. However, data science is more than data analysis, so domain knowledge/theory is important. Data results need to combine with theory-inform interpretation before informing planning and management (Kandt and Batty, 2021; Kang et al., 2019).

The key implications from UDS to support smart transport governance are: 1) conducting robust analysis linking big data and “small data” (Bibri, 2019; Anda et al., 2017; Bettencourt, 2014), 2) putting special focus on citizenry science through mining human-generated data and better exploration of citizens’ needs (Kang et al., 2019; Resch and Szell, 2019), 3) understanding urban dynamics in different time and space scales (Gong et al., 2020; Silva et al., 2021; Batty, 2013), 4) generating interdisciplinary insights through combining data results with domain knowledge (Kandt and Batty, 2021; Kang et al., 2019), 5) linking data with theories to better inform planning and management (Gao and O’Neill, 2020; Landrigan et al., 2018; Benavente-Peces and Ibadah, 2020; Pasichnyi et al., 2019; Szczepanek, 2020; Gong et al., 2020).

2.5 Discussion:

2.5.1 The nexus of CTC and UDS

During the literature review in the previous sections, the nexus between CTC and UDS as well as their linkages with smart transport governance were found.

The nexus of CTC and UDS is the theory-inform data-driven knowledge discovery processes. Firstly, we highlight the important role of theory in interpreting the data results and providing reliable findings to support smart governance. CTC is a set of theories that can be applied in data preparation and result interpretation during the data-driven knowledge discovery. With CTC and domain theories informed, the data-informed results can be used in urban management.

Smart transport systems, as CASs, are unpredictable, so we need to recognise limitations in UDS and should be transparent on how we use the data and how we interpret data results (Kitchin, 2016). Although perfect predictions cannot be made, we can build “good enough” robust models to understand the dynamic urban systems and how they function. Guided by CTC, we can capture the spatiotemporal dynamics in human activities, sentient environments and interactions with the help of new detailed, finer-scale data or advanced algorithms, so that new hypotheses on urban dynamics can be generated (Kandt and Batty, 2021).

Additionally, the latent possibilities of uncertain configurations and future scenarios in smart transport should be recognised. Latent approaches from UDS can be applied to build or improve urban data analytics, models and simulations for decision making and problem solving, handling, and making sense of urban data with a degree of uncertainty. With an adaptive understanding of uncertain urban issues, smart governance can be more elasticities and with higher adaptive capacity.

2.5.2 A holistic framework for smart transport governance

When it comes to smart transport governance, we suggest a citizen-centric adaptive framework for smart city governance, with a special focus on smart transport. The new conceptual framework is based on the existing smart governance framework (see Figure 2-1), the mostly used analytical framework in existing smart transport governance studies (i.e., multi-level perspective of socio-technical transition theory) in section 2.2, implications from CTC in section 2.3, and implications from UDS in section 2.4. We integrated the key notions and implications from previous sections to build a holistic framework for understanding the smart city and transport system. Interdisciplinary approaches are added to the framework, providing adaptive and robust understanding of smart transport governance. The new framework is in Figure 2-3.

The meso-level regime of the smart city and transport system is the dynamic stable practice and structure. The fixity of a regime can be affected by changes in the macro-level landscape such as global trends or changes in micro-level niche emergence such as technological novelties (Geels, 2012; Girones et al., 2019; Manders et al., 2020). Within the smart city regime, the smart city and the six key interrelated subsystems are complex systems with spatiotemporal dynamics. In the six main dimensions of the smart city concept, smart governance overlaps with all other domains and governs the other subsystems. Citizens are put in the centre to avoid technocratic and uneven forms of governance.

The wider macro-level landscape of the smart city regime contains wider trends and crises such as urbanisation, climate change, technological innovations, and pandemics. The micro-level contains the niche contexts that can change the existing regime and shift the current system towards new regimes.

The data-driven knowledge discovery is another main part of the theoretical framework to comprehensively analyse the complex system and potential transitions. Both small and big data can be mined to identify urban dynamics through dynamic data-informed modelling and theory-informed interpretations. Reliable findings can then support the understanding of complex systems and smart governance outcomes.

In the final part of this theoretical framework, the outcomes come with possibilities rather than certain results. We suggest adaptive responses to quick or slow changes and consider alternative governance options. Instead of aiming for certain outcomes of smart governance, uncertainty is added, and the outcomes come with different possibilities. When facing the uncertainty of the ever emergent and changing transport system, we see uncertainties as opportunities and governance can actively make use of the opportunities to ensure transport sustainability, accessibility, innovation, and inclusion.

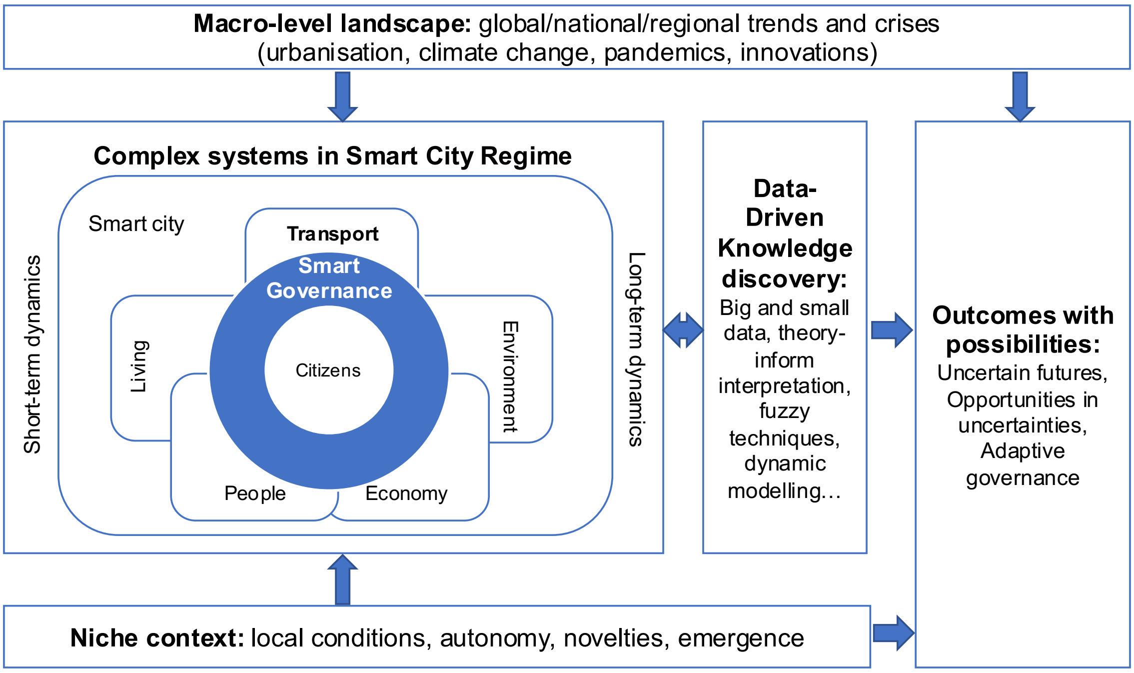

Figure 2-3: Theoretical framework

The theoretical framework can be applied to understand the complex development of smart transport and support decision making. The following chapters are empirical examples of applying the conceptual framework in cases.

The systematic literature review in section 2.2 shows that smart transport has been less assessed and has not been evaluated through a holistic framework. Through the new theoretical framework, we further robustly assess smart transport development in the next chapter.

As citizens are at the centre of the theoretical framework, the correct needs of citizens can be extracted through complex theory-inform and data-driven knowledge discovery in the new framework. New insights on long-lasting transport governance issues such as mode choices, travel demands, interventions' impacts, and short-term changes of niche-level or landscape-level emergence can be found in empirical studies. We further generate new findings of citizens in the following chapters to support a more robust and adaptive understanding of important issues.

Additionally, the unexpected outbreak of the global health crisis from COVID-19 is a disruption from a wider context in the macro-level landscape. Facing this sudden abruption, how it changes transport systems in the short-term and the long-term remains unclear. What are the possible outcomes brought by the pandemic and whether there are opportunities that smart governance can make use of in the highly uncertain post-COVID transport futures are interesting questions that can be addressed through adaptive understanding and governance. The last part of this PhD research will address this gap.

2.6 Conclusion

Smart transport is a key domain in the “smart city” concept and this complex system needs smart governance to ensure the technological advancement brings added public value instead of undermining sustainability and inclusion. Smart transport governance is applying innovations during transport planning and management processes with different stakeholders involved. Recent studies on smart transport governance mainly discussed theoretical notions, analytical approaches, main tasks and interrelated dynamic contexts.

The main tasks include understanding the roles of governance, assessing smart transport, extracting travel demands, improving collaborative governance, and inventing uncertain future mobilities. The dynamic local or wider contexts of smart transport systems are another main topic. Smart transport systems should be analysed in the local and wider contexts. The three types of widely used theories are from 1) technology and innovation studies, 2) political science and public policy studies, and

3) complexity theory. The most used conceptual framework in existing smart transport governance studies is from socio-technical transition theory. Complexity perspectives have been widely seen but seldom explicitly discussed in the reviewed articles.

Regarding the analytical approaches, multidisciplinary methods with various data sources and techniques have been applied to explore the main tasks. The importance of data science has been emphasised in the articles, but existing studies have mainly focused on data analysis.

The chapter then conducted a broader review of CTC and UDS to extract the main implications for supporting smart transport governance. The key implications of CTC are: 1) accepting latent possibilities of uncertain futures, 2) exploiting opportunities in uncertainties, 3) understanding complex issues with interdisciplinary knowledge and approaches, 4) using holistic/robust analysis to support decision making, 5) increasing the responsiveness to changes through adaptive governance, 6) enhancing adaptivity in new institutional frameworks, and 7) adjusting governance to temporal and spatial dynamics.

From UDS, five main guidance are found, which are: 1) conducting robust analysis linking big data and “small data”, 2) putting special focus on citizenry science through mining human-generated data and better exploration of citizens’ needs, 3) understanding urban dynamics in different time and space scales, 4) generating interdisciplinary insights through combining data results with domain knowledge, 5) linking data with theories to better inform planning and governance.

The nexus of CTC and UDS is further identified in the theory-inform data-driven knowledge discovery processes. Transparent and robust analysis that accepts unpredictability and uncertainties as well as adaptive planning based on fuzzy data models are important linkages.

Building on the literature review, we proposed a holistic framework for smart transport governance. The citizen-centric adaptive framework for smart city and transport governance is based on the existing smart governance framework, multi-level perspectives of socio-technical transition theory, and main implications from the CTC and UDS. The proposed framework contains a macro-level landscape, complex systems in a smart city regime, niche context, data-driven knowledge discovery, and outcomes with possibilities.

We made three contributions to existing literature through this chapter. First, we systematically reviewed the recent empirical and conceptual studies on smart transport governance and identified the main research topics. Second, by combining relevant works on CTC and UDS, we draw additional insights to enhance the framework of smart urban and transport governance. Third, we proposed an integrated framework that links key notions and methods to support smart transport governance.

We conclude that the employment of complexity theory and data science can contribute to smart transport governance through adaptive planning and dynamic data-driven knowledge discovery. The following chapters will apply the proposed framework to empirical studies.

Chapter 3 : A comparative analysis of smart transport using the most used indicators in the literature juxtaposed with interventions in English metropolitan areas

3.1 Introduction

The smart transport system is an essential part of the “smart city” concept and operations in a smart city. Technological innovation has permeated this sector for decades, allowing smart transport to be a priority in smart city development. “Smart” in the transport sector can refer to new propulsion (e.g., electricity), new vehicle controls (e.g., Intelligent Transport System), new business models (e.g., car-sharing), and new transport planning and policies. Their main objectives are reducing pollution, reducing traffic congestion, increasing safety, improving transfer speed and reducing travel costs (Benevolo et al., 2016). However, the emergent technologies and the changing travel patterns caused by smart transport innovations are highly uncertain and complex. Advanced technology and proper governance should enhance smart transport development with added values (Docherty et al., 2018). Transport governance can influence the transforming directions of smart transport development. Smart transport governance contains a set of strategies, schemes, policies, projects and actions, including integrated ticketing, travel apps, electric vehicles, automated vehicles, and sustainable transport policies (Woods et al., 2017; Harriss and Kearney, 2021). Thus, it is necessary to evaluate and analyse smart transport developments in terms of technologies, methods, infrastructure, and interventions.

The UK has a long history of developing smart transport by applying new technologies and planning for smart future mobility. For example, the UK Department for Transport

(DfT) developed a pioneering travel information service, Transport Direct, in the early 2000s (DfT, 2017). The UK's overall policy aims to be a world leader in Intelligent Transport Systems (ITS) (DfT, 2017). Transport innovations, including integrated information, cleaner vehicles, and infrastructure, are highlighted in many cities' latest transport plans. Transport authorities such as Transport for London (TfL) are actively preparing for smart future mobility (Government Office for Science, 2019). However, the smart transport development in English cities has not been systematically assessed. Therefore, it is worth investigating the development of smart transport in English cities.

One way of measuring smart transport is through indicator analysis. An indicator or index can represent a specifically evaluable phenomenon through proper measurements (Lopez-Carreiro and Monzon, 2018), making it a powerful tool to describe complex phenomena and support decision-making processes (Kitchin et al., 2015; Battarra et al., 2018a). Many studies and some international standards organisations have proposed indicators to benchmark smart transport as part of the smart city index. Despite many indicators and indices to assess the smart city, less work has been done to evaluate the smart transport system and compare smart transport in different areas. A comprehensive and up-to-date framework with a holistic set of indicators/indices to measure various aspects of the smart transport system is necessary (Anthopoulos et al., 2019; Yousif and Fox, 2018; Battarra et al., 2018a).

The new evaluation framework is built on the previous highly cited research by Debnath et al. (2014), Garau et al. (2016), Pinna et al. (2017a), Lopez-Carreiro and Monzon (2018) as well as Battarra et al. (2018a). We first review articles on smart transport indicators and selected the most used indicators to represent important variables in smart transport. This chapter further supplies the existing list of indicators with new indicators that can reveal trending topics. We then synthesise the individual indicators into three groups to reflect specific aspects and the overall development. Firstly, indicators in the private, public, and emergency transport sub-systems are aggregated into private transport index, public transport index, and emergency transport index, respectively. Secondly, the individual indicators are synthesised into accessibility, (environmental) sustainability, and innovation indices, which are the most common aggregated indices in the literature. Lastly, the accessibility, sustainability, and innovation indices are further combined into a composite index - the smart transport index to show the overall development of smart transport in a city. The indicators, indices and composite index are applied in English metropolitan areas, specifically the Combined Authorities (CAs) and Greater London (GL).

In this research, we first discuss the smart city and transport development in the English metropolitan areas, with a special focus on their governance and interventions. This sub-regional spatial scale is suitable for the analysis of transport networks and the governance structures in the sub-regional tier can have strong impacts on transport development. We then review the most commonly used indicators and indices in analysing smart transport. Based on the review, a new evaluation framework containing important individual indicators and aggregated indices is constructed for our empirical study. The new framework is applied to the eleven selected English metropolises, allowing us to compare the smart transport developments in the current CAs and the GL. This chapter raises three sub-questions:

1) How do the English metropolises govern smart transport in terms of interventions?

2) What are the most common and important indicators and indices to examine smart transport?

3) What are the smart transport developments in the English metropolises?

The next sections are as follows: Section 3.2 presents the smart city and smart transport features in English metropolitan areas; Section 3.3 outlines the method of indicator selection and index construction for this chapter; Section 3.4 discusses smart transport results using our proposed evaluation framework in the selected cases; Section 3.5 extensively discusses the linkages among smart transport development, indices, and interventions as well as the implication for smart transport governance; Section 3.6 concludes the results and points to further research directions.

3.2 Smart city and smart transport in English metropolitan areas

3.2.1 Overview of English metropolitan areas

A metropolitan area, also known as a functional urban area, usually contains at least one urbanised core area with a substantial population and adjacent districts (Moreno-Monroy et al., 2021). The core areas and the surrounding areas are spatially, socially and economically linked (Hall, 2009). Combining the resources in a metropolitan area can accelerate urban development and enhance the local and regional transport systems (Fenwick and Johnston, 2020). This chapter uses the metropolitan scale because the transport network can be comprehensively analysed at this scale. The government authority at this scale can effectively operate economies and transport systems through interventions (National Audit Office, 2017; Marsden and Docherty, 2019).

In England, Combined Authorities (CAs) are mostly built at the metropolitan scale with devolutions of some powers and functions, aiming to accelerate economic development and improve transport networks outside London (Lorencka and

Obrebska, 2018). CAs are built according to *the Local Democracy, Economic Development and Construction Act 2009* and *the Cities and Local Government Devolution Act 2016* (Sandford, 2017). CAs, as legal structures, can coordinate the resources and interventions in their areas. Although different CAs have their priorities. In main strategies, transport is one of the main focuses of all CAs (except North of Tyne). Each devolution deal contained a transport budget for improving transport governance, including bus franchising, local roads, and rail network (Sandford, 2022). The administrative authorities can regulate smart transport providers to ensure the new products are accountable and guarantee added value to the areas. Local and sub-regional authorities can steer their smart transport development towards socially, environmentally and economically sustainable directions by actively inventing mobility futures (Moscholidou and Pangbourne, 2019).

The first CA was Greater Manchester CA in 2011, followed by four CAs (i.e., North-East, Liverpool City Region, Sheffield City Region[1] and West Yorkshire) in 2014 (Sandford, 2019a). Tees Valley and West Midlands were established in 2016. In 2017, West of England and Cambridgeshire and Peterborough CAs were built. North of Tyne was separated from North-East CA in 2018 (Uberoi, 2021). It should be noted that North-East and North of Tyne cooperate in transport governance. The North East joint transport committee makes collective decisions across the region (NECA, 2022). When designing CAs, policymakers followed the concept of “city region”, indicating that CAs have strongly relied on their core cities (Hickman and While, 2017). An exception is Cambridgeshire and Peterborough, which is not a former city region (Hickman and While, 2017). Cambridgeshire and Peterborough CA is the combination of two relatively independent boroughs (Townsend, 2019). Although it is not a typical

[1] Renamed as South Yorkshire in 2021

metropolitan area, this chapter keeps it for comparison. The ten existing CAs at the time of data collection (May 2020) are used as case studies.

When comparing CAs, Greater London is often used as a beacon [e.g. (Sandford, 2019b)]. The Greater London Authority (GLA) can be seen as the first successful “Combined Authority”. It was formed in 2000 (the first mayoral election), covering the areas of 33 local authorities (Hickman and While, 2017; Townsend, 2019). GLA is an example of devolution and the mayor holds powers to promote urban and transport developments within London (Sandford, 2018). TfL, chaired by the mayor, is in charge of delivering transport services such as tubes, buses, trams, taxis and private hire vehicles. It should be noted that the GLA is a unique authority, and it is fundamentally different from the CAs. The Mayor of London and GLA were created by different legislation and were given different powers (Fenwick and Johnston, 2020). The establishment of GLA and TfL was subject to the *Greater London Authority Act 1999*. Additional power was further added by the *Greater London Authority Act 2007* (Sandford, 2022). Nevertheless, Greater London is the largest metropolis in England. Thus, we choose the ten CAs and Greater London as cases of metropolitan areas to illustrate the development of smart transport in metropolitan areas.

The English metropolitan areas are heterogeneous in population, area, and social-economic development (as shown in Table A-1, Appendix A). Greater London has the largest population, population density, and total gross value added (GVA). Greater London is the sole first-tier metropolis. Considering both population and total GVA, the second-tier large metropolises are West Midlands, Greater Manchester, West Yorkshire, and Liverpool City Region. The remaining metropolises either have smaller populations or worse economic performance. The social-demographic information has been considered when constructing the indicators and analysing the results.

3.2.2 Smart city developments in case areas

Smart cities in the UK can be categorised into four groups: leaders, contenders, challengers and followers, according to the latest UK smart city report (Woods et al., 2017). The report evaluated smart cities based on their vision, digital innovation, service innovation, sustainability plans, and stakeholder engagement (Woods et al., 2017). Leaders have clear and inclusive smart city planning as well as projects in full-scale levels. Contenders are cities have significant strategies and projects; however, they are gaps in strategies. Challengers are with a vision for smart city and begin to work on smart city projects. For cities that just begin to build smart city (with limited pilot projects), they are classified as followers (Woods et al., 2017).

The core cities (i.e., primary local authority) in the metropolitan areas are ranked in this report. Bristol in West of England and London are the leaders. Contenders include Manchester in Greater Manchester, Birmingham in West Midlands, Leeds in West Yorkshire, Peterborough and Cambridge in Cambridgeshire and Peterborough, and Newcastle in North of Tyne. Sheffield in Sheffield City Region and Liverpool in Liverpool City Region are challengers (Woods et al., 2017).

In each core city, the main innovation areas vary. Bristol leads on the Internet of Things (IoT) while London is ahead in data and analytics. One of the notable data sources in London is the TfL. Leeds is a model of innovative health, whereas Peterborough is the pioneer in sustainability. The focus of Newcastle is education (Woods et al., 2017).

To simplify the policy reviewing process, the smart interventions in this study refer to policies or projects that use digital technologies or accelerate the deployment of technologies. The ongoing smart city interventions in metropolitan areas are identified and listed in Table 3-1. For each metropolitan area, all interventions listed on their official websites were collected. Policies, strategies, programmes, packages, projects, and schemes were all included.

West of England has six main interventions, focusing on the economy and business, transport, and energy. Great London has an extensive list of smart strategies, the most important of which is “Smart London Together”, which is the Mayor’s roadmap to make London “the smartest city in the world” (Smart London, 2018). London’s interventions cover many areas, The key areas are transport, sustainability, health, energy, economy and business, as well as data and analytics.

Greater Manchester has many projects in its Digital Strategy, with special focuses on economy and business, health, data and analytics, transport, and sustainability (GMCA, 2020). West Midlands has a main industrial strategy and five smart projects, targeting economy and business, transport, the IoT, and sustainability (WMCA, 2019). Eight ongoing smart interventions are found in West Yorkshire, in which principal areas are economy and business, transport and sustainability (WYCA, 2021). Cambridgeshire and Peterborough CA has three initiatives and two proposed smart transport schemes (CPCA, 2021). The key areas are sustainability as well as the economy and business. Two programmes aiming at economy and business are seen in North of Tyne (NoT, 2021).

Concentrating on economy and business, Sheffield City Region has three smart interventions (SYMCA, 2021). Liverpool City Region has a range of action plans related to smart city development, and its key areas cover economy and business, sustainability, and transport (LCR, 2021). Two interventions focusing on the economy and business as well as sustainability are seen in North East CA (NECA, 2021). Tees Valley has several smart city projects, mainly in the areas of economy and business (TVCA, 2021).

Generally, more interventions can be found in areas with a higher-ranking smart core city, as shown in Table 3-1. West of England, Greater London, Greater Manchester, and West Midlands have more policies, schemes, and projects, with more key areas than other CAs. In terms of key areas, economy and business are the most common focus of all authorities, followed by sustainability and transport. Transport is one of the key areas in West of England, Greater London, Greater Manchester, West Midlands, West Yorkshire, and Liverpool City Region.

Table 3-1: Smart city interventions

Authorities	**Ongoing smart interventions**	**Key areas**	**Smart core city ranking and category**
West of England	Regional Public Transport, 5G Smart Tourism Future Bright, Women into Digital Jobs, Education and Training programme, Creative Scale Up, Clean Growth	Economy and Business, Transport, Energy	1 Leader
Greater London	Smart London Together, London Datastore, TfL Open Data Portal, Police Interactive Dashboard, CleanTech, FLexLondon, Energy for Londoners, Smart Sustainable Districts, Smart Mobility Living Lab, etc.	Transport, Sustainability, Energy, Health, Economy and Business, Data and Analytics	2 Leader
Greater Manchester	Digital Strategy (Early Years Digitisation, Integrated Digital Healthcare Record, Greater Manchester Information Sharing Strategy, Smart Ticketing, Greater Manchester Cyber and Resilience, Made Smarter and Digital Enablement Services, ERDP-funded Digital Initiatives, Annual Digital Creative and Tech Festival, etc), Digital Response to COVID-19, Greater Manchester Tech Fund, Smart Energy project (Smart Heat project)	Economy and Business, Health, Data and Analytics, Transport, Sustainability, Energy	3 Contenders
West Midlands	West Midlands Industrial Strategy, Digital Retraining scheme, 5G Testbed, Energy Innovation Zones, Midlands Future Mobility, Mobility Credits Pilot	Economy and Business, Transport, Internet of Things, Sustainability	4 Contenders
West Yorkshire	Growing business (Connecting Innovation), Clean energy and Environmental Resilience (Ultra-Low Emission Vehicle Taxi Scheme), Clean Bus Technology Fund, Energy Accelerator, West Yorkshire-plus Transport Fund, The Leeds Public Transport Investment Programme (LPTIP Real Time Programme), Local Transport Plan and DfT Funding	Economy and Business, Transport, Sustainability	5 Contenders
Cambridgeshire and	Eastern Agri-Tech Growth Initiative, Growth Hub Projects, Digital Sector Strategy	Sustainability, Economy and	9 Contenders 10

Peterborough		Business	Contenders
North of Tyne	STEM and Digital Skills Programme, Inclusive Economy Innovation Fund (Employability and Skills Programme)	Economy and Business	14 Contenders
Sheffield City Region	Knowledge Gateway, AMRC Light weighting Centre, Local Growth Fund (Barnsley's Digital Media Centre)	Economy and Business	16 Challengers
Liverpool City Region	Adult Education Budget (digital skills), Skills for Growth Action Plan (Innovation Action Plan 2018-2020, Digital and Creative Action Plan 2018-2020, Skill Strategy 2018-2023, Health & Care Action Plan 2018-2020, Low Carbon Action Plan 2018-2020, Advanced Manufacturing Action Plan 2018-2020), Strategic Investment Fund (Train Connectivity and Information Systems), Smart Ticketing business case	Economy and Business, Sustainability, Transport	18 Challengers
North East	Strategic Economic Plan (smart specialisation), Go Ultra Low North East	Economy and Business, Sustainability	-
Tees Valley	BoHo "The Digital City", Hartlepool College of Further Education (Telecare and Electric Vehicle Skills Enhancement), Hartlepool Centre of Excellence in Technical Training for the Creative Industries, Inspiring our Future	Economy and Business	-

3.2.3 Smart transport governance and interventions in CAs and GLA

Smart transport aims to enhance accessibility, environmental sustainability, safety and other factors (Giffinger et al., 2007). A smart transport system incorporates such elements as infrastructure, travel means, products, business models, operation systems and interventions. Smart mobility considers the delivery of people, data, and goods. In this chapter, we focus on transporting people within metropolitan areas.

In English metropolitan areas, most of the CAs are responsible for the transport systems and their services. In the devolution deals, all CAs (except North of Tyne) have transport-related powers such as leading the sub-regional transport plan. North of Tyne, which was separated from the North East CA, does not have new transport powers in its devolution deal (HM Government, 2018). The preeminent intervention in North East was published in 2016, before the establishment of North of Tyne, and it covers the areas of the current North of Tyne and North East combined authorities.

Greater London, West Midlands, and Greater Manchester have their unique government bodies responsible for regulating the transport system and coordinating transport services. TfL was created in 2000. It runs the day-to-day transport operations, including buses, the undergrounds, and taxis, and it manages the transport infrastructure in Greater London. It is known as an internationally leading transport body (Government Office for Science, 2019; White, 2016). Following TfL's success, Transport for Greater Manchester was set up in 2011 and Transport for West Midlands was founded in 2016, aiming for "London-style" powers and delivering "London-style" transport (GMCA, 2022; WMCA, 2022).

Apart from the smart transport interventions listed in Table 3-1, we also reviewed the main ongoing transport policies in each case to see if they include smart transport. The mentioned smart aspects and the key objectives of accessibility, (environmental) sustainability, and innovation are listed in Table 3-2.

Each authority has its main transport intervention, and all the documents mentioned some smart transport elements. Most of the transport plans have a separate chapter discussing the new technologies and smart mobility possibilities. Thus, most of the interventions have the main objectives of innovation. Smart ticketing, smart information, and cleaner vehicles and infrastructure (e.g., electric cars and electric charging devices) are the most commonly highlighted smart elements in the main transport interventions. Other smart aspects, including smart logistic delivery, smart parking, open data, CAV and MaaS, are also mentioned in many documents, mainly discussing the potential impacts of coming technologies.

Generally, all transport authorities in the English metropolitan areas admit that smart technologies can influence future mobility and transport systems and they need to prepare for the potential changes. The main challenge is that smart technologies are highly uncertain in terms of their impacts on existing urban structures and transition paths that they may lead to. Although all authorities are preparing for future mobility, transport planners and policymakers cannot plan for the new smart transport products or business models because many future scenarios are possible. Thus, an evaluation framework with good indicators to illustrate the current situations and future potentials of a city’s smart transport can provide meaningful insights. The insights can help decision makers understand the uncertain smart transport in the English metropolitan areas.

Table 3-2: Smart transport interventions

Authorities	Transport authorities	Main transport interventions	Smart aspects	Main Objectives
Greater London (GL)	Transport for London	Mayor's Transport Strategy 2018	Communication, Smart Information, Smart Ticketing, CAV, Sharing Services, Smart Logistics Delivery, Cleaner Vehicles and Infrastructure, Open Data, Smart Parking	Accessibility, Sustainability, Innovation
West Midlands (WM)	Transport for West Midlands	Movement for Growth: The West Midlands Strategic Transport Plan	Smart Information Systems, Maas, Open Data, Clean Air Zone, Intelligent Traffic Management, Smart Logistics Delivery, Smart Road Safety, Sharing Service, Smart Motorways, Smart Ticketing, CAV	Accessibility, Sustainability, Innovation
West of England (WE)	WECA	Joint Local Transport Plan 2020-2036	CAV, Maas, Open Data, Smart Motorway, Intelligent Traffic Management, Smart Information, Cleaner Vehicles and Infrastructure, V2I [2] Communication, Smart Logistics Delivery, Smart Ticketing, Sharing Services	Accessibility, Sustainability, Innovation
Greater Manchester (GM)	Transport for Greater Manchester	GM Transport Strategy 2040	Maas, Smart Information, Smart Ticketing, Cleaner Vehicles and Infrastructure, Sharing Services, Smart Vehicles	Accessibility, Sustainability, Innovation

[2] Vehicle to infrastructure

			(CAV), Smart Signal, Smart Traffic Control, Smart Motorway, Open Data	
Liverpool City Region (LCR)	LCRCA	LCRCA transport plan	Smart Ticketing, On-Demand Bus Service, Cleaner Vehicles and Infrastructure	Accessibility, Sustainability
North of Tyne (NT)	NTCA[3] & NECA	-	-	-
West Yorkshire (WY)	WYCA	Transport Strategy 2040	Smart Ticketing, Smart Motorway, Smart Information, Open Data, Intelligent Traffic Management, Maas, CAV, Sharing Services	Accessibility, Sustainability, Innovation
Cambridgeshire and Peterborough (CP)	CPCA	Local Transport Plan	Smart Motorway, Smart Information, Smart Infrastructure, Cleaner Technology, Smart Parking	Accessibility, Sustainability
North East (NE)	NECA	Transport Manifesto	Smart Ticketing, Smart Information, Cleaner Vehicles and Infrastructure, Intelligent Traffic Management, Sharing Services	Accessibility, Sustainability
Sheffield City Region (SCR)	SCRCA	SCR Transport Strategy	Maas, Smart Ticketing, Smart Motorway, CAV, Smart Logistics Delivery, Smart Information	Accessibility, Sustainability
Tees Valley (TV)	TVCA	Strategic Transport Plan	Cleaner Vehicles and Infrastructure, Smart Information, Smart Ticketing, Maas, CAV	Accessibility, Sustainability

[3] No new transport power in the devolution deal. A new transport committee working with NECA.

3.3. Methodology

Indicators and composite indices have been used in many studies to measure the performance of complex systems such as the urban transport system (Giffinger et al., 2007; Battarra et al., 2018a; Battarra et al., 2018b; Debnath et al., 2014; Kitchin et al., 2015). An indicator is a measuring instrument to describe an element, process, or property of a system, representing any level of complexity. An index can be seen as complex indicators to capture complex interrelations or conditions, showing a system level complexity (Heink and Kowarik, 2010). This chapter used a four-step method to construct the index system, addressing methodological gaps 2. We first reviewed the international standard organisations for smart city and transport documents and scientific studies in Web of Science, Scopus and Google Scholar that used indicators or indices to evaluate smart transport. Secondly, we built an appropriate set of indicators for our case studies based on the systematic literature review. Thirdly, six synthetic indices were calculated from the selected indicators in each subset. To illustrate different sub-systems, we then aggregated the indicators into public, private and emergency transport indices. Categorised by important pillars in smart transport, indicators were also aggregated into accessibility index, sustainability index, innovation index and smart transport index. Finally, this chapter presented a composite index, namely the smart transport index.

3.3.1 Measuring smart transport through indicators and indices

The keywords of “smart transport/ transportation/ mobility” and “index/ indicator” were applied when searching for the relevant documents in the International Organisation for Standardization and academic databases of Web of Science, Scopus, and Google Scholar. The search returned 301 articles, and a bibliometric analysis was conducted (Aria and Cuccurullo, 2017). Irrelevant documents, including articles that do not contain indicator/index, studies focusing on other smart features (e.g., smart environment), measurements only on a specific aspect of smart mobility (e.g., road maintenance, walkability) and full-text articles unavailable to access online, were removed in the screening process. After screening, 39 publications were left, and these articles were used for choosing indicators.

In the reviewed studies, more than 50 different indicators have been used to describe various aspects of the smart transport system, covering the main themes of accessibility, service, safety, technological integration, and equity. We identified the 30 most used indicators, as listed in Table 3-3. The indicators that have been used more than ten times are low-emission vehicle (17), public transport supply/service (16), integrated and electronic ticketing system (15), cycling lane (14), bike-sharing (14), mode choice (14), car-sharing (13), modern parking solution (12), traffic coordination/operation system (11), and real-time travel planner (10). Indicators with fewer than two citations were not included.

In the reviewed papers, most indicators are one of two types: measurable indicators such as the number of vehicles, and “on and off” indicators such as whether a city has travel ticketing online (e.g., 0 for no travel ticketing online and 1 for a city with online ticketing). In calculating the index, most authors used the rescaling method in normalisation (e.g., Min-Max normalisation) and equal weighting in aggregating individual indicators. Most of the indices are calculated by the geometric or arithmetic means of different variables.

Authors have classified indicators into several subsets. Each subset indicates an important aspect of smart transport. Indicators in each subset are often aggregated into an index to illustrate an aspect. In classifying indicators, Debnath et al. (2014) analysed smart transport in three categories (i.e., private transport, public transport and emergency transport). Battarra et al. (2018a) used ICT, sustainability, and accessibility variables to evaluate smart mobility. The main sub-themes in the existing literature are private transport system, public transport system, innovation (mainly ICT-based) and sustainability. For example, all sustainability-related indicators can be combined into the sustainability index (Battarra et al., 2018a). In Table 3-3, the identified indicators can be grouped into three themes: accessibility, (environmental) sustainability, and innovation (Battarra et al., 2018a; Pop and Prostean, 2019).

Diverse groups' key themes and indicators assess smart transport in numerous studies. However, the categories in some studies could not represent the whole picture of the smart transport system in a city. For example, emergency transport is often neglected in smart transport research. Many indicators are related to the sharing economy, but many studies only use bike-sharing and car-sharing as important innovation indicators. Other advanced innovations, such as MaaS, are not included in the existing index. Regarding empirical studies, the most studied cases are Italian and Spanish cities. UK cities have not yet been thoroughly assessed. Thus, a topology of smart transport development in UK cities can contribute to the existing literature.

Table 3-3: Most used indicators in reviewed articles

Indicators	Sources	Themes
Public transport supply/service	(Battarra et al., 2018a, Battarra et al., 2018b, Das, 2020, Garau et al., 2015, Garau et al., 2016, Indrawati et al., 2017, Li et al., 2019, Lopez-Carreiro and Monzon, 2018, Miguel et al., 2018, Ogrodnik, 2020, Petrova-Antonova and Ilieva, 2018, Pinna et al., 2017, Lerner et al., 2011, Pop and Proștean, 2019, Shaheen et al., 2019, ISO 37122:2019;)	Accessibility
Low-emission vehicle	(Bakogiannis et al., 2019, Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Das, 2020, Indrawati et al., 2017, Lopez-Carreiro and Monzon, 2018, Miguel et al., 2018, Mol, 2018, Petrova-Antonova and Ilieva, 2018, Pinna et al., 2017, Yigitcanlar et al., 2020, Zapolskyte et al., 2020, Zong et al., 2019, Pop and Proștean, 2019, Shaheen et al., 2019, ISO 37122:2019;)	Sustainability
Integrated and electronic ticketing system	(Aleta et al., 2017, Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Debnath et al., 2014, Garau et al., 2015, Garau et al., 2016, Longo et al., 2019, Lopez-Carreiro and Monzon, 2018, Petrova-Antonova and Ilieva, 2018, Pindarwati and Wijayanto, 2015, Zapolskyte et al., 2020, Zong et al., 2019, Pop and Proștean, 2019, ISO 37122:2019)	Innovation
Cycling lane	(Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Garau et al., 2015, Garau et al., 2016, Miguel et al., 2018, Mol, 2018, Ogrodnik, 2020, Orlowski and Romanowska, 2021, Petrova-Antonova and Ilieva, 2018, Pinna et al., 2017, Zapolskyte et al., 2020, Pop and Proștean, 2019, ISO 37120:2018)	Sustainability, Accessibility
Bike-sharing	(Balducci and Ferrara, 2018, Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Braga et al., 2019, Garau et al., 2015, Garau et al., 2016, Mol, 2018, Petrova-Antonova and Ilieva, 2018, Pinna et al., 2017, Zapolskyte et al., 2020, Lerner et al., 2011, Pop and Proștean, 2019, ISO 37122:2019)	Sustainability, Innovation
Car-sharing	(Balducci and Ferrara, 2018, Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Garau et al., 2015, Garau et al., 2016, Li et al., 2019, Mol, 2018,	Innovation, Accessibility

	Petrova-Antonova and Ilieva, 2018, Zapolskyte et al., 2020, Lerner et al., 2011, Pop and Proștean, 2019, ISO 37122:2019)	
Mode choice	(Das, 2020, Dudzevičiūtė et al., 2017, Indrawati et al., 2017, Li et al., 2019, Lopez-Carreiro and Monzon, 2018, Mol, 2018, Orlowski and Romanowska, 2021, Petrova-Antonova and Ilieva, 2018, Yigitcanlar et al., 2020, Zong et al., 2019, Lerner et al., 2011, Pop and Proștean, 2019, Shaheen et al., 2019, ISO 37120:2018)	Accessibility
Modern parking solution	(Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Debnath et al., 2014, Garau et al., 2015, Garau et al., 2016, Mol, 2018, Pinna et al., 2017, Wibowo and Grandhi, 2015, Zapolskyte et al., 2020, Pop and Proștean, 2019, ISO 37122:2019)	Innovation
Traffic coordination/operation system	(Aleta et al., 2017, Benevolo et al., 2016, Debnath et al., 2014, Indrawati et al., 2017, Lopez-Carreiro and Monzon, 2018, Mol, 2018, Orlowski and Romanowska, 2021, Pindarwati and Wijayanto, 2015, Wibowo and Grandhi, 2015, Zapolskyte et al., 2020, Pop and Proștean, 2019)	Innovation
Real time travel planner	(Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Debnath et al., 2014, Garau et al., 2015, Garau et al., 2016, Mol, 2018, Petrova-Antonova and Ilieva, 2018, Pindarwati and Wijayanto, 2015, Zapolskyte et al., 2020)	Innovation
Travel time	(Abu-Rayash and Dincer, 2020, Indrawati et al., 2017, Longo et al., 2019, Lopez-Carreiro and Monzon, 2018, Miguel et al., 2018, Petrova-Antonova and Ilieva, 2018, Lerner et al., 2011, Shaheen et al., 2019, ISO 37120:2018)	Accessibility
Restricted/special traffic zone	(Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Debnath et al., 2014, Petrova-Antonova and Ilieva, 2018, Pindarwati and Wijayanto, 2015, Wibowo and Grandhi, 2015, Zapolskyte et al., 2020, Pop and Proștean, 2019)	Sustainability
Intelligent traffic light/ Smart street lighting	(Balducci and Ferrara, 2018, Battarra et al., 2018a, Battarra et al., 2018b, Garau et al., 2015, Garau et al., 2016, Mol, 2018, Zapolskyte et al., 2020, ISO	Innovation

	37122:2019)	
Mobile phone app	(Aleta et al., 2017, Battarra et al., 2018a, Battarra et al., 2018b, Das, 2020, Garau et al., 2015, Garau et al., 2016, Liu et al., 2019, Orlowski and Romanowska, 2021, Pop and Proștean, 2019)	Innovation
Public transport demand	(Battarra et al., 2018a, Battarra et al., 2018b, Garau et al., 2015, Garau et al., 2016, Li et al., 2019, Lopez-Carreiro and Monzon, 2018, Petrova-Antonova and Ilieva, 2018, Pinna et al., 2017, Pop and Proștean, 2019)	Accessibility
Variable message sign	(Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Garau et al., 2015, Garau et al., 2016, Zapolskyte et al., 2020, Pop and Proștean, 2019)	Innovation
In-vehicle technologies: AVL, CCTV, detection, GPS	(Battarra et al., 2018a, Battarra et al., 2018b, Debnath et al., 2014, Mol, 2018, Pindarwati and Wijayanto, 2015, Zapolskyte et al., 2020, ISO 37122:2019)	Innovation
Pedestrian zone	(Balducci and Ferrara, 2018, Battarra et al., 2018a, Battarra et al., 2018b, Benevolo et al., 2016, Mol, 2018, Zapolskyte et al., 2020, Pop and Proștean, 2019)	Sustainability, Accessibility
Road fatality rate	(Bakogiannis et al., 2019, Das, 2020, Indrawati et al., 2017, Lopez-Carreiro and Monzon, 2018, Ogrodnik, 2020, Shaheen et al., 2019, ISO 37120:2018)	Sustainability
Private transport supply	(Indrawati et al., 2017, Ogrodnik, 2020, Orlowski and Romanowska, 2021, Salvia et al., 2016, Lerner et al., 2011, ISO 37120:2018)	Accessibility
Autonomous vehicle	(Benevolo et al., 2016, Kelley et al., 2020, Mol, 2018, Zapolskyte et al., 2020, ISO 37122:2019)	Innovation
Sustainable mobility plans/measures/investment	(Aleta et al., 2017, Indrawati et al., 2017, Lopez-Carreiro and Monzon, 2018, Orlowski and Romanowska, 2021, Zapolskyte et al., 2020)	Sustainability
Electronic bus stop sign	(Battarra et al., 2018a, Battarra et al., 2018b, Garau et al., 2015, Garau et al., 2016, Zong et al., 2019, Pop and Proștean, 2019)	Innovation
Electric charging device	(Abu-Rayash and Dincer, 2020, Benevolo et al., 2016, Mol, 2018, Petrova-Antonova and Ilieva, 2018, Zapolskyte et al., 2020)	Sustainability

Mobility difficulty	(Indrawati et al., 2017, Lopez-Carreiro and Monzon, 2018, Miguel et al., 2018, Shaheen et al., 2019)	Accessibility
Internet access/service	(Das, 2020, Liu et al., 2019, Orlowski and Romanowska, 2021, Petrova-Antonova and Ilieva, 2018)	Innovation
Park and ride	(Balducci and Ferrara, 2018, Ogrodnik, 2020, Zapolskyte et al., 2020)	Innovation
Air quality	(Bakogiannis et al., 2019, Lerner et al., 2011, Shaheen et al., 2019)	Sustainability
Road transport energy consumption	(Indrawati et al., 2017, Mol, 2018)	Sustainability
Travel cost	(Indrawati et al., 2017, Lopez-Carreiro and Monzon, 2018)	Accessibility

3.3.2 Building a smart transport evaluation framework for English metropolitan areas

The new evaluation framework was built on the review above and followed a typical indicator selection process (Sdoukopoulos et al., 2019). We first selected the most used indicators from the literature and supplied existing indicators with potential new indicators. We finalised individual indicators by checking quality selection criteria, data availability and duplication. To illustrate the latest trends in smart transport, we added five new indicators to the final indicator list. After building the individual indicator list for evaluating the specific aspect of smart transport technologies, methods and infrastructures, the indices were constructed by scaling up indicators from different subsets. Six indices (i.e., private transport index, public transport index, emergency transport index, accessibility index, sustainability index and innovation index) were aggregated. Finally, a composite index was calculated to quantify the overall development of smart transport. A comprehensive and detailed picture of smart transport in each case can be demonstrated using the individual indicators. The key elements in smart transport (e.g., the overview of the public transport system) can be shown by aggregated indices (e.g., public transport index). We can further compare the general situations in various cases using the composite index. The individual indicators, aggregated indices and composite index make up the new evaluation framework for smart transport.

3.3.2.1 Building a disaggregated indicator list

The selection of smart transport indicators in this study followed a four-step procedure. Firstly, the systematic review in the last section allowed us to identify the most common indicators. The most used indicators (in Table 3-3) were included in our thorough list of potential indicators. Some of the indicators can be illustrated by several detailed indicators. For example, public transport supply/service can be represented by bus/

rail/ metro length/ network, depending on the data availability in each study.

The second step constructed the potential new indicators to supplement the current indicators. Policy documents, reports and articles have also discussed other new themes that have not been included in the current indicators. Private-hire cars, shared travel, Mobility-as-a-Service (MaaS, i.e., one-stop online intermodal journey planner), intersections between physical and digital infrastructures, data and connectivity, electrification, decarbonisation, automation, and new business models are trending themes in governmental documents (Government Office for Science, 2019). MaaS, Internet of Things (IoT) and open data have been mentioned in the future mobility chapters of many transport interventions. Academic studies from recent years also discuss IoT (Mohanty et al., 2016; Mohammadian and Rezaie, 2020; Crainic et al., 2019; Shaheen et al., 2019), open data and data-driven products (Shaheen et al., 2019; Tomaszewska and Florea, 2018; Kumar et al., 2018; Xu and McArdle, 2018), MaaS (Cruz et al., 2018; Anthony et al., 2020; Li et al., 2019a; Finger and Audouin, 2019), self-driving or driverless cars (Šurdonja et al., 2020; Toh et al., 2020) and emergency service tracking (Šurdonja et al., 2020).

Among these new themes, MaaS, IoT, self-driving vehicles and open data are not included in the existing indicator set. New innovation indicators on these four themes can supplement the current set. Additionally, indicators of emergency transport systems are rare in the existing literature. Indicators for normal ambulance performance and performance in a time of pandemic as well as smart ambulances are added to illustrate the accessibility and innovation of a smart emergency transport system.

Thirdly, we checked the quality selection criteria and data availability for all potential indicators to finalise indicators. The criterion for selecting indicators contains measurability, ease of availability, interpretability and the isolability of transport impact (Castillo and Pitfield, 2010). Regarding data, the sources in this study include the National Travel Survey 2017 (NTS) (DfT, 2020a), road accidents and safety statistics, vehicle statistics, bus statistics, rail statistics from the DfT (DfT, 2020b), Highways England (Highways England, 2020), Centre for Connected and Autonomous Vehicles (CCAV) (Centre for Connected and Autonomous Vehicles, 2020), as well as statistics from the Department for Business, Energy and Industrial Strategy (BEIS) (Department for Business, 2020), Office for National Statistics (ONS) (Office for National Statistics, 2020), National Health Service (NHS) (National Health Service, 2020), and Public Health England (PHE) (Public Health England, 2020). The main dataset in this study is NTS 2017. We mainly used data from 2017 to 2020 to illustrate the latest smart transport situations in English metropolitan areas.

Simultaneously with the previous work, we also reviewed webpages of city authorities, services providers, related companies, consulting reports and news media, using Google Search Engine to collect information. We searched for keywords and city names (Debnath et al., 2014; Pindarwati and Wijayanto, 2015). Eleven indicators used the data from webpages through searching in Google. Apart from restricted schemes and trial CAV projects, other indicators are on/off to reduce the potential error of miscounting. For new indicators, we had data only on MaaS and Open data Application Programming Interface (API) in public transport, as well as ambulance disposition rate, ambulance disposition rate changes due to the pandemic, and connected ambulance in emergency transport. Thus, the five new indicators were added to supply the existing toolkit.

Fourthly, as the indicators were equally weighted in aggregating into indices, we also checked the duplication of similar indicators and delete similar indicators to avoid over-representing one aspect. For example, the mobility difficulties (e.g., personal disability and poor connections) in private transport can be illustrated by either the number of blue badges (disabled parking permits) or the percentage of users who have mobility difficulties in cars. We chose the latter one as it contains the most comprehensive information.

Based on the four-step procedure, the final evaluation framework contains 38 existing indicators and five new indicators (broad in Table 3-4, 3-5, 3-6, column 2). We chose three subsets to show different transport sub-systems in a city, namely private transport (including walking and cycling), public transport and emergency transport systems. In each subsystem, we further classified the indicators into three themes: accessibility, sustainability, and innovation. Accessibility concerns “the ability of places to be reached”, relevant resources (e.g., car for private transport and bus for public transport) and affordable costs for local people (Battarra et al., 2018a). Sustainability considers environmental aspects such as energy consumption, and social-economic aspects include issues such as road safety. It should be noted that sustainability index in this chapter only measures environmental sustainability. Innovation deals with new technologies and new business models used in the transport system.

Three types of indicators are listed: 1) percentage indicators (N%) such as private vehicle rate (PV_A2_vehiclerate); 2) number indicators (N) such as the number of urban access regulation schemes (PV_S1_restrictedschemes); and 3) on and off indicators (1/0) such as whether a metropolitan area has CAV hard infrastructure (PV_I9_CAVhardinf). Ideally, percentage and number indicators should be used to show detailed information in each field. When detailed information is not available or accessible, we use the on and off indicators. Binary indicators are mostly about innovation features, showing the presence or absence of each innovative product or service. The data sources we accessed cannot provide more accurate information (e.g., the actual number/percentage) in these features, so we used the on and off value. The types of indicators are shown in the unit column.

The indicators can be either positive or negative. Positive indicators mean the indicators have a positive impact on the corresponding theme. For example, car access with cars or vans in a household leads to greater accessibility, making PV_A3_caraccess a positive indicator. On the contrary, negative indicators are the factors that can decrease the level of each theme. For instance, the higher mean of particulate matter (PM2.5) indicates less environmental sustainability. The detailed indicators of each variable are listed below (Table 3-4, 3-5, 3-6).

Table 3-4: Selected indicators for private transport

Themes	**Indicators**	**Description of indicators**	**Unit**	**P/N**	**Data sources**
Accessibility	PV_A0_traveltime	Average minimum travel time to reach the nearest key services by car, 2017	N	+	DfT
	PV_A1_roadnetwork	Road network length, Km/km2	N%	+	DfT
	PV_A2_vehiclerate	Number of private vehicles per inhabitants	N%	+	DfT
	PV_A3_caraccess	car access (with car/ van)	N	+	NTS
	PV_A4_modechoice	% Private modes (car and van, motorcycle, other private transport)	N%	+	NTS
	PV_A5_mobdifficulties	% Mobilities difficulties in cars	N%	-	NTS
	CW_A1_modechoice	% Walking and cycling mode	N%	+	NTS
	CW_A2_footdifficulties	%Mobilities difficulties on foot	N%	-	NTS
Sustainability	PV_S1_restrictedschemes	urban access regulation schemes: low emission zones, urban road tolls, other access regulation	N	+	Google
	PV_S2_ecologicalcars	% Ultra-low emission vehicles (ULEVs) licensed in all registered vehicles, 2019	N%	+	DfT
	PV_S3_electriccharging	Publicly available electric vehicle charging devices per 100,000 inhabitants by local authority, 2019	N%	+	DfT
	PV_S4_airquality	Population-weighted annual mean of pm2.5, 2018	N	-	UK Air
				-	
	PV_S5_roadfatalityrate	Number of road fatalities by car per 100,000 inhabitants, 2019	N%	-	PHE
	PV_S6_roadenergyconsumption	Road transport energy consumption (Tonnes of oil	N	-	BEIS

		equivalent of diesel, petrol cars and motorcycles), 2017			
	CW_S1_roadfatalityrate	Number of road fatalities by walking and cycling per 100,000 inhabitants, 2019	N%	-	DfT
Innovation	PV_I1_carclub	Car-sharing demand, number of car club members per 1000 inhabitant	N%	+	NTS
	PV_I2_PHV	Ride-sourcing supply, licensed private hire vehicles per 1000 inhabitant	N%	+	DfT
	PV_I3_smartmotorway	% Number of operational smart motorways in total road length	N%	+	Highway England
	PV_I4_mobilealarm	SMS/ mobile notification for traffic alert	1/0	+	Google
	PV_I5_VMS	Variable message sign/ matrix sign/ Variable Signs and Signals	1/0	+	Google
	PV_I6_realtimeforcast	Real-time traffic forecast	1/0	+	Google
	PV_I7_internetaccess	% Internet users, 2019	N%	+	ONS
	PV_I8_CAVsoftinf	CAV soft infrastructures: virtual labs	1/0	+	CCAV
	PV_I9_CAVhardinf	CAV hard infrastructures: testbeds	1/0	+	CCAV
	PV_I10_CAVproject	Number of trial CAV projects	N	+	Google
	PV_I11_ITSproject	Intelligent Transport System projects funded by DfT	1/0	+	DfT

Table 3-5: Selected indicators for public transport

Variables	Indicators	Description of indicators	Unit	P/N	Data sources
Accessibility	PB_A0_traveltime	Average minimum travel time to reach the nearest key services by public transport and walking, 2017	N	+	DfT
	PB_A1_busservice	Public transport supply: Vehicle kilometres on local bus services by local authority, 2018/2019 (million)	N	+	DfT
	PB_A2_busjourney	Public transport demand: passenger journeys on local bus services 2018/19 (million)	N	+	DfT
	PB_A3_modechoice	% Public mode (bus, London underground, rail, taxi, other public transport)	N%	+	NTS
	PB_A4_buscosts	Local bus fares index (at current prices) by metropolitan area status and country	N	-	DfT
	PB_A5_busdifficulties	% Bus difficulties	N%	-	NTS
Sustainability	PB_S1_ecobus	ecological buses?	1/0	+	Google
	PB_S2_energyconsumption	Road transport energy consumption (Tonnes of oil equivalent of diesel, petrol cars and motorcycles), 2017	N	-	DfT
	PB_S3_roadfatalityrate	Number of road fatalities per 100,000 inhabitants, 2019	N%	-	DfT
	PB_S4_interventions	Sustainable actions/ objectives in public transport in smart transport intervention?	1/0	+	Policy review
Innovation	PB_I1_CCTV	% Buses used as Public Service Vehicles with CCTV by metropolitan area status and country, local bus operators only	N	+	DfT

	PB_I2_AVL	% Buses used as Public Service Vehicles with automatic vehicle location (AVL) device by metropolitan area status, local bus operators only	N	+	DfT
	PB_I3_AVLrealtimeifo	% Buses with an AVL to provide real-time service information to customers by metropolitan area status, local bus operators only	N	+	DfT
	PB_I4_Wifi	% Buses used as Public Service Vehicles with free Wi-Fi by metropolitan area status and country, local bus operators only	N	+	DfT
	PB_I5_MaaS	Mobility as a service?	1/0	+	Google
	PB_I6_buslane	detection of unauthorised vehicles: Have a bus lane/ bus only/ bus gate enforcement system?	1/0	+	Google
	PB_I7_contactlessticket	% Buses with live EMV readers that can accept contactless payment cards1 by metropolitan area status, local bus operators only	N%	+	DfT
	PB_I8_integratedticket	% Buses with live readers that accept Oyster/ ITSO Smart-cards1 by metropolitan area status, local bus operators only	N%	+	DfT
	PB_I9_openapi	Open data platform/ API?	1/0	+	Google

Table 3-6: Selected indicators for emergency transport

Variables	Indicator	Description of indicator	Unit	P/N	Data source
Accessibility	ET_A1_ambulancedispositionrate	% Number of emergency ambulance dispositions/ inhabitant, in May 2019	N%	+	NHS
	ET_A2_pandemicchange_ambulancedispositionrate	% Increase in number of emergency ambulance dispositions compared to normal time, pandemic period (March 2020) and normal time (May 2019)	N%	+	NHS
Innovation	ET_I1_signals	Emergency vehicle priority signal – able to provide priority signal?	1/0	+	Google
	ET_I2_connectedambulance	Trial digital/ connected/ smart ambulance? Ambulance Global Digital Exemplars?	1/0	+	Google

3.3.2.2 Synthesizing the smart transport indicators into aggregated indices

The units and results of the selected indicators in the previous section vary. To build an index by aggregating individual indicators, the result of each indicator needs to be rescaled into a common range. We applied the most used method in existing studies, namely Min-Max normalisation, to rescale the results (Garau et al., 2016; Garau et al., 2015; Lopez-Carreiro and Monzon, 2018). In all positive indicators except PV_A0_traveltime and PB_A0_traveltime, a larger number indicates better performance. For other positive indicators, the best result will be rescaled to 1 and the worst to 0, using the formula (1). Negative indicators will be rescaled from 0 (the worst) to 1 (the best), using the formula (2). For PV_A0_traveltime and PB_A0_traveltime, less time used in travelling to the key services means better accessibility. Thus, the rescaling of two travel time indicators is special, using the formula (2).

Positive indicators:

$$Xir = \frac{Xi - Min(Xi)}{Max(Xi) - Min(Xi)} \quad (1)$$

Negative and special indicators:

$$Xir = \frac{Xi - Max(Xi)}{Min(Xi) - Max(Xi)} \quad (2)$$

Where:

Xir: re-scale value of X_i

Xi: initial score of the indicator

$Min(Xi)$: the minimum value of the indicator

$Max(Xi)$: the maximum value of the indicator

On and off indicators do not need to be recalculated. The rescaling process was done in R (Team, 2013). After rescaling, the indicators can be aggregated into indices.

Three groups of indices were constructed: 1) three synthetic indices for different transport systems, 2) three synthetic indices for different themes (i.e., Accessibility, Sustainability, and Innovation), and 3) smart transport index. Following the commonly used synthetic approach (Battarra et al., 2018a; Battarra et al., 2018b; Lopez-Carreiro and Monzon, 2018), we calculated the average value of all indicators in each category (Formula 4) to construct the synthetic indices. Previous studies have weighted indicators or sub-systems to reveal the relative importance of different elements in smart transport systems (Lopez-Carreiro and Monzon, 2018; Li et al., 2019b).

Expert/stakeholder opinions have often been used to decide the weights. This method requires extensive time and resources to collect data (Debnath et al., 2014). Thus, we decided to use equal weight for all variables and individual indicators in this study, as most authors did (Garau et al., 2016; Garau et al., 2015; Lopez-Carreiro and Monzon, 2018; Pinna et al., 2017b).

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3}$$

For the first group, private transport index (I_{PV}), public transport index (I_{PB}) and emergency transport index (I_{ET}) were calculated by the mean of indicators in each transport system. The formulas are as below:

$$I_{\mathrm{PV}} = \frac{\sum_{i=1}^{n} PVi}{n} \tag{4}$$

$$I_{\mathrm{PB}} = \frac{\sum_{i=1}^{n} PBi}{n} \tag{5}$$

$$I_{\mathrm{ET}} = \frac{\sum_{i=1}^{n} ETi}{n} \tag{6}$$

Where:

PV: Private transport (including walking and cycling) indicators in Table 3-4

PB: Public transport indicators in Table 3-5

ET: Emergency transport indicators in Table 3-6

Similarly, accessibility index (I_A), sustainability index (I_S) and innovation index (I_I) were calculated in the formulas (7), (8), and (9). These three indices are the most common used aggregated indices to show the three pillars of the transport system.

$$I_{\mathrm{A}} = \frac{\sum_{i=1}^{n} Ai}{n} \tag{7}$$

$$I_{\mathrm{S}} = \frac{\sum_{i=1}^{n} Si}{n} \tag{8}$$

$$I_{\mathrm{I}} = \frac{\sum_{i=1}^{n} Ii}{n} \tag{9}$$

Where:

Ai: Accessibility indicators listed in Tables 3-4, 3-5, 3-6

Si: Sustainability indicators listed in Tables 3-4, 3-5, 3-6

Ii: Innovation indicators listed in Tables 3-4, 3-5, 3-6

3.3.2.3 Composing the smart transport index

Finally, the smart transport index (I_{ST}) was defined through the formula below, which merges the three main dimensions (i.e., accessibility, sustainability, and innovation) in an area. It is difficult to decide which of the three dimensions is most important in the smart transport evaluation framework. Weights for each factor may vary from case to case. For example, stakeholders in London may give a different weight to the sustainability index from stakeholders from Great Manchester. Weighting variables for English cities may not be transferrable to cities in other countries. Thus, in constructing the I_{ST}, the three dimensions of accessibility, sustainability, and innovation, are equally weighted in the composite index. This study used the geometric mean (Formula 10) to show the overall transport performance (Garau et al., 2016; Garau et al., 2015; Lopez-Carreiro and Monzon, 2018; Pinna et al., 2017b).

$$I_{\mathrm{ST}} = \sqrt[3]{I_{\mathrm{A}} \times I_{\mathrm{S}} \times I_{\mathrm{I}}} \tag{10}$$

3.4. Results

This section contains the results of six aggregated indices and the composite index. In the first part, we present the results for three main transport systems, namely private transport, emergency transport and public transport in the eleven cases. In the second part, a comparison of smart transport in terms of accessibility, sustainability and innovation is provided. The last part shows the result for the smart transport index.

3.4.1 Private, public, and emergency transport indices

The value of the private transport index varies in each case (see Figure 3-1). The best private transport system is obtained in Greater London (0.710); followed by West of England (0.637) and West Midlands (0.628), Liverpool City Region (0.551), and Greater Manchester (0.545). Generally, the metropolis with a greater population and better economic performance has a better private transport system. Higher GVA may lead to more resources available for local government, which can be allocated to transport infrastructures (e.g., road network) and intelligent transport projects. For example, Greater London and West Midlands have sustainable schemes such as low-emission zones and more connected and automated vehicles (CAV) test infrastructures. Additionally, people in wealthier places are more likely to have access to private cars. The top five metropolises in private transport also have higher vehicular densities. It should be noted that cycling and walking mode is counted in private transport. Relevant indicators on the non-motorised mode indicate that walking and cycling difficulties occur in many road networks in metropolises. Roughly speaking, the places with better social-demographic background have worse performance in cycling and walking.

For the public transport index, the best performance is also obtained in Greater London (0.735), followed by West Midlands (0.608), Liverpool City Region (0.603), Greater Manchester (0.593) as well as Cambridgeshire and Peterborough (0.587). These places with better economic performance provide more bus services and experience fewer bus difficulties. These metropolises have a wide range of smart technologies in their public transport system. For example, most of their buses are equipped with closed-circuit television (CCTV), automatic vehicle location (AVL) and smart ticketing systems. They all have open data platforms for developers to make use of real-time and high-volume transport data for improving their transport applications and services. These applications can benefit public transport users in these places. Also, the top areas have pilot projects or plans for MaaS, which is seen as a future user-centric trend in public transport. The MaaS is believed to benefit public transport and active mode (GO-Science, 2019). Relevant projects are City Mapper and London Transport Planner in Greater London, iMove in Greater Manchester, Whim in West Midlands, CAPITALS in Liverpool City Region, and the intelligent City Platform for Cambridgeshire and Peterborough (Bevis, 2018).

For the emergency transport index, North of Tyne (0.873), Tees Valley (0.873), and North East (0.873) rank first among all cases, followed by West Midlands (0.700). The North East Region working with North East Ambulance Service NHS Foundation Trust (including North of Tyne, Tees Valley and North East CAs) has the best performance in terms of both accessibility of ambulance service and smart technology used in ambulances. The value of emergency transport does not correspond to the social-demographic information.

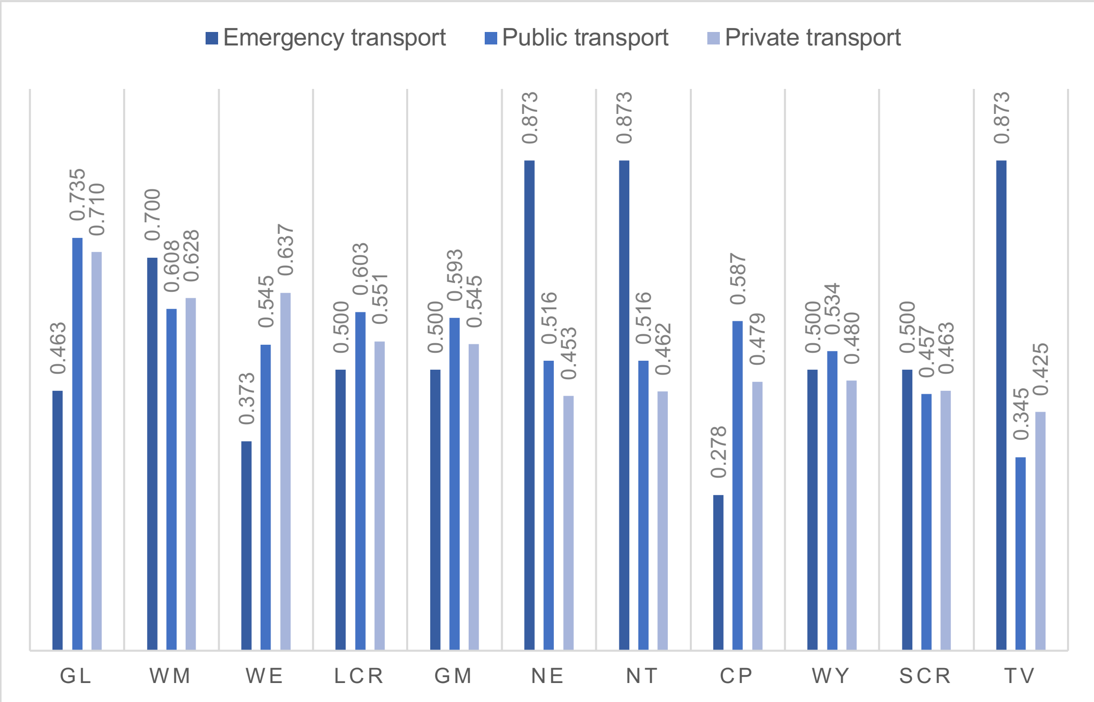

Figure 3-1: Results of different transport systems

3.4.2 Accessibility, sustainability, innovation indices

For the accessibility index, the values range from 0.297 in Sheffield City Region to 0.614 in Greater London. The most accessible metropolitan areas are Greater London (0.614), West of England (0.477), West Midlands (0.460), Liverpool City Region (0.418) and Tees Valley (0.405). Greater London has the best accessibility in its public transport and private transport (including cycling and walking) in all areas. West of England and West Midlands show particularly good accessibility in all transport systems. People can access to good private, public, and non-motorised transport in Liverpool City Region. People in Tees Valley can access good emergency, public and private transport.

For the sustainability index, the best performances are obtained by North of Tyne (0.730), Tees Valley (0.703), and West of England (0.674). The worst area is Greater

London (0.456). West of England has the most sustainable public transport system and a very sustainable private transport system, while Greater London has the most sustainable private transport system and the least sustainable public transport. North of Tyne and Tees Valley have high performance in terms of environmental sustainability in public and private transport systems. Although all the areas have ecological buses or plan to introduce ecological buses, areas with more bus services witness more energy consumption by public transport. For private transport, London has the greatest number of schemes to manage air pollution, as well as the most ecological vehicles and electric charging devices. However, it also has the highest road energy consumption and the worst air pollution.

Values of the innovation index range from 0.313 to 0.883 in different English metropolitan areas. The most innovative areas are Greater London (0.883), West Midlands (0.768), Greater Manchester (0.673), Liverpool City Region (0.656) and West of England (0.608). Greater London has the best innovative capacities in both public and private transport systems. Smart technologies have been used in London’s transport system, including CCTV, AVL devices, smart tickets, pilot MaaS and open data in public transport, as well as car-sharing services, intelligent transport system projects, CAV infrastructures and projects in private transport. West Midlands ranks second in its innovation in both private and public transport. It has free Wi-Fi, MaaS, integrated tickets, and open data in its public transport system. It aims to build smart future mobility, and it is now one of the premier CAV testbeds. 5G is also used to improve the connected transport system in West Midlands. Liverpool City Region and Greater Manchester have excellent innovative public transport, with smart devices, ticketing systems and pilot MaaS. West of England is another important testbed for CAV projects, so it also has innovative private transport.

However, the most innovative areas are not the places ranked highest in emergency transport innovation. Innovation in emergency transport includes emergency vehicle priority signals and trail connected ambulance projects. South Central Ambulance Service NHS Foundation Trust, West Midlands Ambulance Service NHS Foundation Trust and North East Ambulance Service NHS Foundation Trust are working on the digitally advanced ambulance to become Ambulance Global Digital Exemplars (National Health Service England, 2020a), which refers to “an internationally recognised NHS provider delivering improvements in the quality of care, through the world-class use of digital technologies and information” (National Health Service England, 2020b). Thus, the most innovative places in emergency transport are North of Tyne, North East and Tees Valley.

3.4.3 Smart transport index

The smart transport index (I_{ST}) considers the dimensions of accessibility, sustainability, and innovation in transport systems, which are the three main pillars of smart transport. The index is a tool to summarise and simplify the overall smart transport developments with multidimensions in each case. Ranking the results of the I_{ST} can compare the divergent performances in smart transport in the selected cases.

The result shows that Greater London (0.628) is the smartest among the eleven metropolitan areas, with the best accessibility and innovation performance. The other top smart transport areas are West Midlands (0.591), West of England (0.580), Liverpool City Region (0.549), and Greater Manchester (0.533). The ranking is listed in Table 3-7. As shown in the map (Figure 3-2), the northern areas have worse performances than the southern cities. Generally, Greater London in the first-tier metropolis ranks the first in smart transport index. Those in the second-tier metropolises have high rankings in the smart transport index. One exception is West of England, with a relatively small population, but a very high ranking in terms of its smart transport index.

Table 3-7: Smart transport results for metropolitan areas

Areas	Accessibility index (I_A)	Sustainability index (I_S)	Innovation index (I_I)	Smart transport Index (I_{ST})	Ranking
Greater London	0.614	0.456	0.883	0.628	1
West Midlands	0.460	0.585	0.768	0.591	2
West of England	0.477	0.674	0.608	0.580	3
Liverpool City Region	0.418	0.605	0.656	0.549	4
Great Manchester	0.357	0.629	0.673	0.533	5
North East	0.366	0.617	0.565	0.504	6
North of Tyne	0.299	0.730	0.567	0.498	7
Cambridgeshire and Peterborough	0.335	0.627	0.566	0.492	8
West Yorkshire	0.338	0.581	0.583	0.486	9
Sheffield City Region	0.297	0.599	0.517	0.451	10
Tees Valley	0.405	0.703	0.313	0.447	11

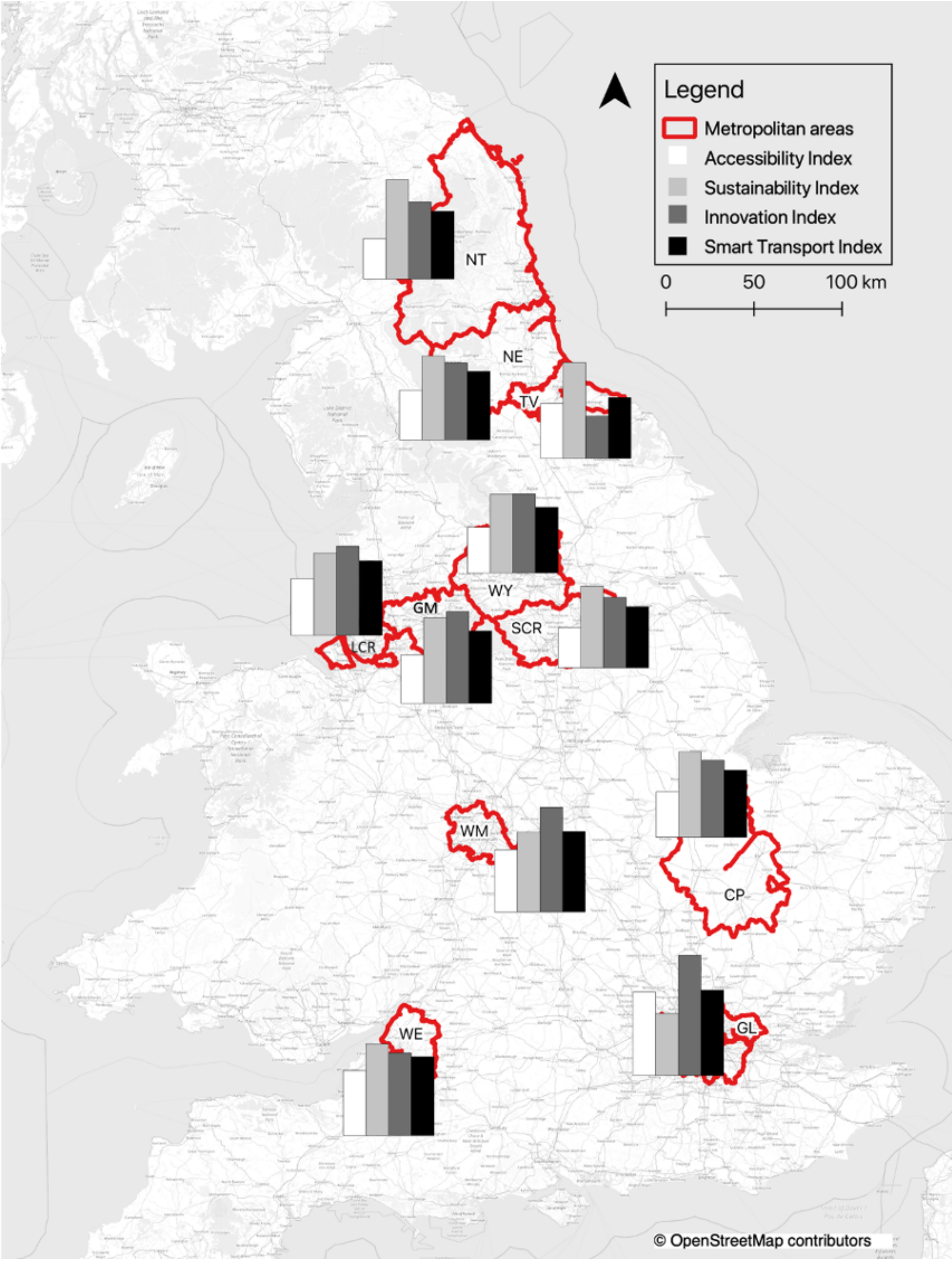

Figure 3-2: Smart transport indices in English metropolitan areas

3.5. Discussion

3.5.1 Linkages between indices and interventions

We ranked the eleven English metropolitan areas by the six aggregated indices that show the main aspects of the transport system and compared the rankings of each index with the composite smart transport index. As shown in Figure 3-3, the areas with the highest rankings in the overall smart transport systems generally have good rankings in private and public sub-systems. The high-ranking areas usually have good accessibility and environmental sustainability. However, good rankings in emergency transport are often not seen in the top areas. These top-ranking areas often have low scores in sustainability.

Considering the social-demographic status of metropolises, the sole first-tier city Greater London ranks first in smart transport. The second-tier cities generally have better scores than the third-tier areas. The two exceptions are West Yorkshire and West of England. The relatively poor accessibility in West Yorkshire prevents it from having a high smart transport ranking. On the contrary, West of England is relatively small and thinly populated, but it has excellent accessibility and sustainability, especially in its private transport system. The social-demographic status of metropolises may be positively linked to the smart transport ranking.

Smart transport is a key component of a smart city. In the UK, the leaders in smart city development (i.e., Greater London and West of England) also have top rankings in smart transport, and probably all other main sectors. While other smart cities each have a different innovation focus, as mentioned in Section 2.2, the ranking of smart transport in the contender group varies (see columns 1 and 2, Table 3-8). A smart city in the contender group may have a less-smart transport sector.

Roughly speaking, the regions that have transport as one of the key areas in their smart city focus and highlight innovative objectives in their main transport interventions (see columns 5 and 6, in Table 3-8) are likely to rank high in the smart transport index. This indicates that political attention is likely to be positive to the development of smart transport. Areas that set transport as a political focus rank higher than other areas in smart transport index, except West Yorkshire. In the cases where innovation is not the main objective of their main transport strategies, the smart ranking tends to be low, as illustrated in North East, Sheffield City Region and Tees Valley. An exception is Liverpool City Region. Innovation is not a key goal or a main chapter in its transport plan, although this plan mentions smart ticketing and smart motorways. Liverpool City Region has several smart transport projects and a relatively good accessibility index. Another exception is West Yorkshire, as explained above. Initiatives and political attention could have a positive impact on the development of smart transport in a city.

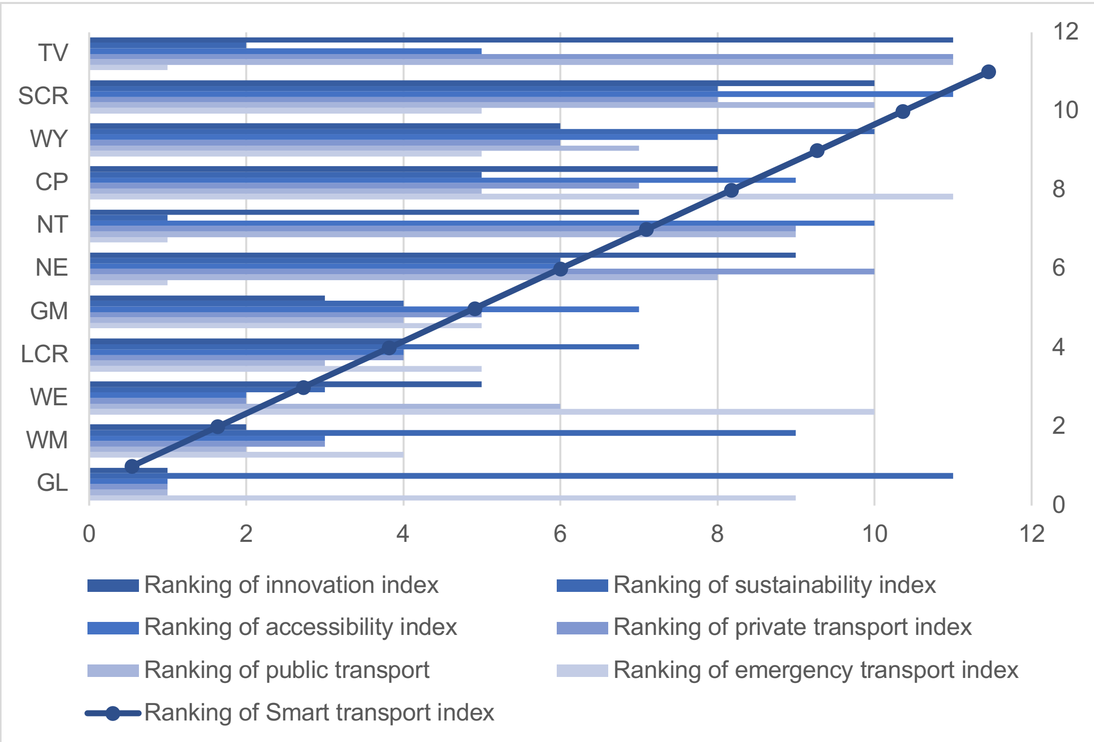

Figure 3-3: Comparison of index ranking

Table 3-8: Comparison between smart transport ranking and interventions

Cases	Smart transport ranking	Smart city category	Social-demographic status (tier)	Transport as key area in Smart city policies	Main Objectives in transport plans
Greater London	1	Leaders	1	yes	Accessibility, Sustainability, Innovation
West Midlands	2	Contenders	2	yes	Accessibility, Sustainability, Innovation
West of England	3	Leaders	3	yes	Accessibility, Sustainability, Innovation
Liverpool City Region	4	Challengers	2	yes	Accessibility, Sustainability
Great Manchester	5	Contenders	2	yes	Accessibility, Sustainability, Innovation
North East	6	-	3	no	Accessibility, Sustainability
North of Tyne	7	Contenders	3	no	-
Cambridgeshire and Peterborough	8	Contenders	3	no	Accessibility, Sustainability
West Yorkshire	9	Contenders	2	yes	Accessibility, Sustainability, Innovation
Sheffield City Region	10	Challengers	3	no	Accessibility, Sustainability
Tees Valley	11	-	3	no	Accessibility, Sustainability

3.5.2 Adaptive governance for smart transport futures

The rapid technological innovations and fast adoption of novelties in the transport sector are challenges for constructing a smart transport index and managing uncertain transport futures. This requires new variables to be added in time and transport planning to respond to the uncertain futures. Uncertainties have been concerns for policymakers and planners for a long time. As explained in Chapter 2, adaptive planning can help manage the uncertainties and emerging innovation in cities (Rauws, 2017; Sengupta et al., 2016). Facing the emerging innovations, adaptive governance should consider various future scenarios with different triggers and pathways. Adaptive approaches from CTC, including simulating potential technological changes, reassessing the current plans when facing new triggers, grasping opportunities during changes, creating conditions towards sustainable transition paths, preparing quick and slow responses, and planning for various transport scenarios, can be used to support smart transport governance (Walker et al., 2019; Lyons and Davidson, 2016). These adaptive approaches can increase the flexibility in planning and management, encouraging planners to think through "what-if" scenarios, exploring alternative developing situations, preparing for different pathways, emphasising the planning processes of "becoming", and supporting collaborative governance (Lyons and Davidson, 2016; Chen and Silva, forthcoming).

As shown in Section 3.2, emerging technologies have been highlighted in almost all transport plans in the eleven cases and all transport strategies stated that preparing for mobility futures is needed. Nevertheless, none of the transport plans has a detailed strategy or a certain plan to deal with emerging technologies because there could be many potential future scenarios. Thus, facing highly uncertain future mobility, transport planning needs to increase its flexibility, by using the adaptive governance framework.

3.6. Conclusion

In a time of complexity, the uncertainty of future mobility and ambitious on-going visions of smart transport interventions require robust methodological tools. Tools such as a comprehensive evaluation framework can unveil the full picture of the different dimensions of smart transport development. In this context, we proposed a new framework for evaluating smart transport in English metropolises, based on the most used indicators, current trends, and data availability. This chapter has contributed to the existing literature and current toolkits on smart transport analysis by identifying the most used indicators, constructing a new evaluation framework with multidimensions included, and applying it in new empirical cases. This empirical study has addressed two gaps we identified in the last chapter, which are methodological gap 2 (lacking sophisticated approaches to understanding complexity in dynamic transport systems) and the governance gap 4 (lacking ability to manage and plan for uncertainties).

The evaluation framework in this study contains 44 commonly used indicators and five new indicators. The new indicators are MaaS, Open data API, ambulance disposition rate, ambulance disposition rate changes due to the pandemic, and connected ambulance. The new indicators can gauge smart transport products, services and quality in public transport and emergency transport. The 49 individual indicators were aggregated into three groups of indices. The first group consists of private, public, and emergency transport sub-systems. Most of the previous studies neglect emergency transport, mainly because of data limitations. As it is an important transport sub-system in cities, especially following the recent global health crisis, we were able to include several indicators to illustrate its quality and innovations. The second group contains accessibility, (environmental) sustainability, and innovation indices. The innovation index includes three new indicators, showing emerging technological innovations such as CAV, MaaS, and IoT. The last group is the composite index of the smart transport index. The three pillars of smart transport - namely accessibility, sustainability, and innovation - are aggregated into the final index. The smart transport index can reveal the overall development of the smart transport sector in each city.

The index analysis has also contributed to the existing understanding of smart transport in the UK in terms of overall development, sub-systems, and interventions. The developments in the selected cases vary. The strengths and weaknesses in each case are shown. The findings show that Greater London has the best performance in many sub-systems and aspects, followed by West Midlands and West of England, while the other areas have strengths in various aspects. Zooming into different sub-systems, Greater London, West of England, and West Midlands score best in private transport. The best performance in public transport can be found in Greater London, West Midlands, and Liverpool City Region. The worst private and public transport systems can be found in Tees Valley. Emergency transport results are slightly different from the other systems. North of Tyne, North East, and Tees Valley have the highest rankings while Cambridgeshire and Peterborough has the worst emergency transport As for different aspects of smart transport, the most accessible areas are Greater London, West of England, and West Midlands. The least accessible area is Sheffield City Region. North of Tyne, Tees Valley and West of England are the most sustainable areas. Greater London is the less environmentally sustainable area. The most innovative areas are Greater London, West Midlands, and Greater Manchester while the least innovative place is Tees Valley. As Greater London is the smartest city in the transport development, the following empirical chapters particularly focus on Greater London.

All English metropolitan areas have adopted smart city interventions, with special emphasis on the economy, business, health, transport, and other sectors. Transport is one of the key systems that a smart city can work on. Metropolises with a focus on smart transport tend to score higher in the smart transport index. This means that political attention could be positive for smart transport development. The results or implementation of the policies may vary, but the places with more political attention to smart transport have a better performance in terms of accessibility, sustainability, and innovation. Accessibility and sustainability are common objectives in all transport plans of English metropolitan areas. All transport strategies mentioned emerging technology and stated that there is a need to prepare for future mobility. However, not all authorities listed innovation as the main goal in their transport plans. Although all metropolises realised the importance of preparing for smarter future mobility, most of the transport plans only discussed the possibilities without detailed strategies. This is because future mobility is highly uncertain in terms of emergent technology. To manage uncertain futures, transport governance needs to be shifted from static to adaptive.

These findings have provided useful insights for sub-regional authorities and their transport authorities. Firstly, the results reveal the overall smart transport development in each metropolis and the performances in the sub-systems (i.e., private, public, and emergency) and main aspects (i.e., accessibility, sustainability, and innovation). Each metropolis has its advantages and weaknesses in specific areas; thus, priority areas to be improved can be easily identified. Secondly, potential factors that can influence the development of smart transport include social-demographic background, geographic locations, and interventions. Against the background of the North-South divide and disparities among metropolitan areas, southern and wealthier areas often have the most resources for developing their smart transport. Balancing smart transport development also requires a more even urban development, which is also one reason to build combined authorities outside London. Lastly, as for the interventions in each authority, adaptive transport planning that accepts uncertainties and considers different transition paths. Also, the authorities could actively invent the future of smart mobility through adaptive approaches to support smart city development.

Our proposed framework with multidimensional indicators is used to evaluate the English metropolitan areas in this study. It can also be applied in other spatial tiers in the UK, including local authority level, regional level and even country level, as well as in other countries. Using the evaluation framework in this study, researchers or practitioners can compare the smart transport developments holistically or in detailed dimensions, main subsystems, and key aspects. Adding more cities and areas for comparison using the proposed framework is a further direction.

Smart transport is developing dynamically as niche-level innovations emerge and influence the existing smart transport regime to some extent. This chapter has considered some niche-level innovations and includes the emergency transport system in our evaluation framework. As time goes by, the proposed framework can be further extended by adding new indicators to match future mobility trends and needs. This is also an area for further research.

The research in this chapter is not without limitations. Data availability is a limitation in selecting indicators and constructing the index. Ideally, all variables should be in the same period, such as 2019. Because of the data limitation, we expanded the time period to three years (2017-2020) in this case study. The index shows the result for the most recent three years. Additionally, individual indicators such as bike-sharing are not included in the index because the data are not available or accessible on the metropolitan scale. Further studies can incorporate new datasets on topics such as IoT, 5G and self-driving when new data on these innovations become available. Additionally, the indicators and variables are equal-weighted and have not been validated. Further development of the evaluation framework might use validated and weighted indices. Through soliciting relevant stakeholders' opinions, a weight to each indicator can be introduced and pilot results can be validated and corrected. Another limitation of this study is that we focus mainly on smart technologies in smart cities and smart governance. Further research could broaden the criteria of smart city intervention and provide a more comprehensive review of smart city and transport interventions in the English metropolises. With a deeper understanding of the smart city and smart transport, we can more effectively link indices with interventions.

Chapter 4 : Understanding daily activity-travel sequences of Londoners

4.1 Introduction

One of the important tasks in smart transport literature is to understand activity and travel patterns, as shown in Chapter 2. The activity-travel studies can provide valuable insights into transport systems, including travel demands of individuals, travel behaviours of different groups, and changing patterns caused by niche-level innovations and landscape-level dynamics. The knowledge can then support activity-based transport models, guide transport market segmentation, enhance evidence-based decision making, and smarten management of complex issues (Cho et al., 2019). As the transport system is interrelated with all other urban sub-systems (as mentioned in Section 2.2), a robust understanding of activity-travel patterns can also provide implications for relevant fields such as smart environment, disease spread control, and tourism planning (Xu and Kwan, 2020).

Activity-travel behaviours usually contain structural information (e.g., transport mode, activity type and socio-demographic attributes) and sequential information (i.e., the order relationship among activities or trips) (Xianyu et al., 2017; Moiseeva et al., 2014). An activity-travel sequence jointly considers trip chains and time uses, including information on types and numbers of trips, transitions between activities, ordering among activities, and durations (McBride et al., 2020; Park et al., 2018). A daily activity-travel sequence is an individual's activity during a day, representing the function, purpose, and trip mode in each time interval. Location information is sometimes added to form a complete space-time sequence of an individual in a day (Moiseeva et al., 2014; Cho et al., 2019).

Previous activity-travel studies have analysed characteristics of activity-travel patterns such as variability and complexity (Moiseeva et al., 2014; Dharmowijoyo et al., 2016; Raux et al., 2016), classification of similar patterns and profiles of segmentations (Park et al., 2020; McBride et al., 2020; Song et al., 2021), and activity-based simulations (Kim, 2018; Koushik et al., 2020). With the advancement of UDS, new algorithms have been applied to further investigate the hidden relationships behind travel behaviours, explanatory determinants of activity-travel patterns and the complexity of travel behaviours (Song et al., 2021; Hafezi et al., 2019). This chapter contributes to existing activity-travel studies by providing an adaptive understanding of activity-travel patterns and employing advanced sequential data mining techniques, to further address the methodological gap 2 (i.e., lacking sophisticated approaches to understanding complexity in dynamic transport systems).

Existing studies identified activity-travel sequences in different areas, including Chicago (Xu and Kwan, 2020), California (Allahviranloo et al., 2017), and Canada (Hafezi et al., 2018b). However, few studies have investigated the daily activity-travel sequences in London and other UK cities. As Greater London ranks first in smart transport development in Chapter 3, this chapter uses Greater London as the case area. Previous studies on Londoners' travel-activity patterns have either focused on a sub-group such as public transport users (Langlois et al., 2016; Gkiotsalitis and Stathopoulos, 2020) or a specific behaviour such as mode choice behaviour (TfL, 2017; Langlois et al., 2016), and cycling behaviour (Feng et al., 2020). The sequential activity-travel patterns in London have not been investigated. Thus, this study intends to reveal representative sequences of the main groups of Londoners.

This chapter employs the London Travel Demand Survey (LTDS), an important dataset shared by TfL, for empirical models. We first develop a methodological framework to analyse the daily spatiotemporal patterns of Londoners. Then, the impacts of the emergent COVID-19 pandemic in the early stage (i.e., January to March 2020) are assessed. In this chapter, the early impact of the pandemic refers to the influence on January to March in 2020, which were mostly self-organising behaviours of citizens as strong restrictions (i.e., national lockdown) were put forward in late March (24th March 2020).

Against the backgrounds above, this chapter aims to answer the following sub-questions:

1) How to identify representative patterns of daily activity-travel sequences?

2) What explains the complexity of Londoners' daily activity-travel patterns based on their travel sequences and socio-demographic profiles?

3) What is the difference in Londoners' daily activity-travel sequences before and during the early stage of the pandemic?

To answer these sub-questions, the rest of this chapter is organised as follows. A brief review of related literature is presented in the next section. The study area and the empirical data used in this chapter are introduced in section 4.3, followed by a detailed methodology in the same section. Then, section 4.4 analyses the results of the research. Section 4.5 further discusses the results, limitations, and governance implications. The final section concludes the chapter.

4.2 Literature review

4.2.1 Activity-travel sequence

Activity-travel patterns are the complex outcomes that result from a sequence of inter-dependent decisions based on the needs and constraints of individuals (Daisy, 2018). From UDS, the activities/locations are seen as a collection of chronologically ordered records that contain individual identifiers, spatial and temporal information (Zhang et al., 2019). Activity-travel data usually consists of structural information and sequential information. Structural information contains socio-demographic characteristics, transport modes, and activity types. Sequential information concerns the order relationship among activities or trips (Xianyu et al., 2017; Moiseeva et al., 2014). Structural information has been extensively researched while sequential information has not been properly captured until the introduction of the sequence alignment method (SAM) by Wilson (1998b) in travel behaviour analysis (Xianyu et al., 2017).

Understanding activity-travel sequences is crucial in urban and transport studies. The main methodologies for analysing activity-travel sequences are sequence comparison, clustering, profiling and prediction (Park et al., 2020). To compare sequences, the SAM method from molecular biology has been widely adopted (Shou and Di, 2018). It measures the distance between two sequences by calculating the minimal cost to transform one into the other (Kim, 2014; Wilson, 1998a).

A daily activity-travel sequence refers to an individual's activities in 24 hours, including function, purpose, trip mode, and location in each time interval. A 24-hour time period is commonly used in modelling individuals' activity-travel schedules (Ben-Akiva and Abou-Zeid, 2013). Many activity-based models, including TASHA by Miller and Roorda (2003), CUSTOM by Habib (2018), and the SALT model by Hafezi et al. (2021), have used a 24-hour time frame. Within 24 hours, the activity sequences can be incomplete tours (without start and end at home) (Ahmed et al., 2020b). The 24-hour sequence in activity-travel studies can be treated as either continuous-time or discrete-time episodes. For discrete-time sequences, 5-minute, 15-minute and 30-minute intervals are widely seen in previous studies (Saneinejad and Roorda, 2009; Kim, 2018; Dharmowijoyo et al., 2017; Hafezi et al., 2019). In this analysis, we use a discrete 24-hour time frame to understand patterns in the case study.

Many studies have clustered and classified activity-travel sequences (i.e., exploring activity-travel patterns through clustering and classification methods) (Park et al., 2020; Saadi et al., 2016). Different clustering methods have been applied to find the groups of similar sequences. Broadly speaking, hard and soft clustering are two main clustering approaches. Hard clustering assigns each item to one cluster while soft clustering can assign an object to different clusters with different membership degrees (Ferraro and Giordani, 2020). For example, Crawford et al. (2018) applied the most widely-used hard clustering methods (i.e., Ward's and K-means clustering) and Allahviranloo et al. (2017) used other hard clustering methods (Affinity Propagation and K-medoids methods) to groups activity-travel patterns. Hafezi et al. (2017) used a fuzzy C-means algorithm to find representative patterns of time-use activity. When dealing with complex data, soft clustering methods are increasingly popular because of their nonlinear nature and flexibility in grouping data (Ferraro and Giordani, 2020; Bolin et al., 2014; Li and Lewis, 2016). Four types of soft clustering methods are fuzzy, possibilistic, rough and model-based approaches, among which fuzzy clustering is the most known method (Ferraro and Giordani, 2020). This study intends to use fuzzy clustering methods that capture more flexible information to generate robust and adaptive understanding.

For profiling clustering results, descriptive analysis, corresponding analysis, regression models, structural equation models, and machine learning models have been used in existing studies (Ahmed et al., 2020b; Dumbliauskas and Grigonis, 2020; Dharmowijoyo et al., 2017; Cheng et al., 2019). This chapter deploys the widely used analytical methods to provide detailed characteristics of similar activity-travel patterns.

4.2.2 Clusters and explanatory variables of activity-travel patterns

Existing empirical studies have clustered activity-travel patterns in different case areas. The clusters mainly contain worker, non-worker, and student groups (Allahviranloo and Aissaoui, 2019; Hafezi et al., 2019). For example, Jiang et al. (2012) found eight clusters of patterns from 2007 to 2008 in Chicago, which were stay-at-home groups, students, three worker groups (early-bird, afternoon, and regular), and three adventurer groups (morning, afternoon, and overnight). In Halifax, twelve groups of weekday patterns in 2008 were found by Hafezi et al. (2018a). Six worker groups (7am-3pm worker, 8am-4pm worker, 9am-5pm worker, extended worker, shorter worker, evening worker) and six non-worker groups (students, morning shop, midday activity, afternoon shop, evening activity and stay-at-home clusters) were identified. In California, Allahviranloo et al. (2017) found eight clusters of activity-travel patterns from 2000 to 2001 and twelve clusters from 2010 to 2011.

The literature on explanatory variables of activity-travel patterns has highlighted several types of determinants, including the socio-demographic status of an individual or household, health status, mobility status, built-environment attributes, and time-space constraints (Hafezi et al., 2018b; Jiang et al., 2012). For example, Allahviranloo et al. (2017) selected age, gender, household size, household vehicles, household structure, income, and the flexibility of work as main variables to infer activity-travel patterns. Dharmowijoyo et al. (2017) explored the relationship between socio-demographic variables at the individual level, household characteristics, travel characteristics, built-environment variables, accessibility variables and activity-travel patterns. Hafezi et al. (2018b) used gender, age, education, occupation, the flexibility of schedule, income, dwelling type, relevant license, household vehicles, usual mode, health status and prior activities to predict daily activity-travel patterns. Age, employment status, income and gender are selected as influence factors by Chen et al. (2019). Among these explanatory variables, socio-demographic information has drawn extensive attention because these characteristics can help transport service providers to understand the transport market and assist transport planners to design efficient policies (Chen et al., 2019; Allahviranloo et al., 2017; Prieto et al., 2017).

As for London, Gkiotsalitis and Stathopoulos (2020) classified the travel patterns of smart card and social media users into six types. Liu and Cheng (2017) identified eleven temporal patterns of smart card users. TfL (2017) classified nine groups of Londoners: affordable transitions, city living, detached retirement, educational advantage, family challenge, settled suburbia, students and graduates, suburban moderation, and urban mobility. The seven key explanatory variables used by TfL were propensity to change travel, mode choice, life stage, income, ethnicity, behavioural changes caused by health status, and the use of mobile phones. The previous studies in London were mainly based on trips rather than sequential activities. The patterns of daily activity-travel sequences in London have not been investigated. Thus, this study can contribute to the empirical understanding of Londoners' activity-travel patterns.

4.2.3 Impacts of COVID-19 on activity-travel patterns

Since the end of 2019, the unexpected spread of coronavirus has significantly influenced the transport sector (Gkiotsalitis and Cats, 2021; Goulias, 2021). During the pandemic, private transport has been favoured due to the higher risk of COVID-19 transmission in public transport (Buhat et al., 2020; Pase et al., 2020; Dingil and Esztergar-Kiss, 2021; Cho and Park, 2021; Bari et al., 2021). Alternative to the mass use of public transport and cars, travelling by active transport (i.e. walking or cycling) has become popular (Lovelace et al., 2020; Laker, 2020).

In terms of travel activities, urban mobility before the pandemic has been replaced by "virtual mobility" (Mouratidis and Papagiannakis, 2021). Out-of-home activities have decreased due to technologies, restrictions, and transport services (Mouratidis and Papagiannakis, 2021; Mouratidis et al., 2021). Existing studies in other countries have shown that people reduced recreational trips and commuting trips but increased pick up and drop off activities to/from schools (Aaditya and Rahul, 2021; Anke et al., 2021; Bari et al., 2021). In the UK, Google data suggested that people slightly reduced work trips in Feb 2020. Starting from March 2020, severe drops in recreational activities and commuting activities can be seen. Shopping activities in grocery and pharmacy stores increased by around 15% and then decreased after the lockdown (Google, 2020).

Existing empirical studies found that socio-demographic characteristics, health status, psychological factors (e.g., anxieties and fear) and policy restrictions could have impacted behaviour changes during the pandemic (Dingil and Esztergar-Kiss, 2021; Goulias, 2021; Musselwhite et al., 2021). Within socio-demographic characteristics, age, gender, education, marital status, work type, income level, household size and car ownership can influence mobility changes (Habib et al., 2021; Jiao and Azimian, 2021; Goulias, 2021). In terms of policy responses, the UK did not introduce stringent restrictions (such as lockdown) that can strongly impact travel behaviours at the early stage of the pandemic (Narlikar and Sottilotta, 2021). Thus, the early behavioural changes in the UK can show the adaptiveness of citizens in responding to the emergency. We intend to explore the changes in daily activity-travel sequences from January to March 2020 to show the early impact of COVID-19.

4.3 Methodology

4.3.1 Research framework

To close the methodological gaps 2 and 3 as well as the governance gap 4 of lacking ability to manage and plan for uncertainties , this chapter aims: 1) to identify activity-travel patterns of Londoners, 2) to find main socio-demographic covariates in each activity-travel cluster, 3) to present the early influence of COVID-19 on activity-travel patterns.

As shown in Figure 4-1, the analytical framework of this chapter began with building and identifying daily sequences of individuals from the LTDS dataset. Considering both activities, trips, and location information, we combined activity-trip sequences and location sequences into two-channel sequences. The multi-dimensional sequences were then grouped using clustering methods. To unfold the complexity in each group, we extracted the representative sequences, calculated complex indices, compared activity-trip time use, presented the main mode choice, and found the relationships between patterns and sociodemographic factors. Lastly, policy implications were briefly discussed in the discussion. The framework can also provide detailed insights for policies to be explored in further research.

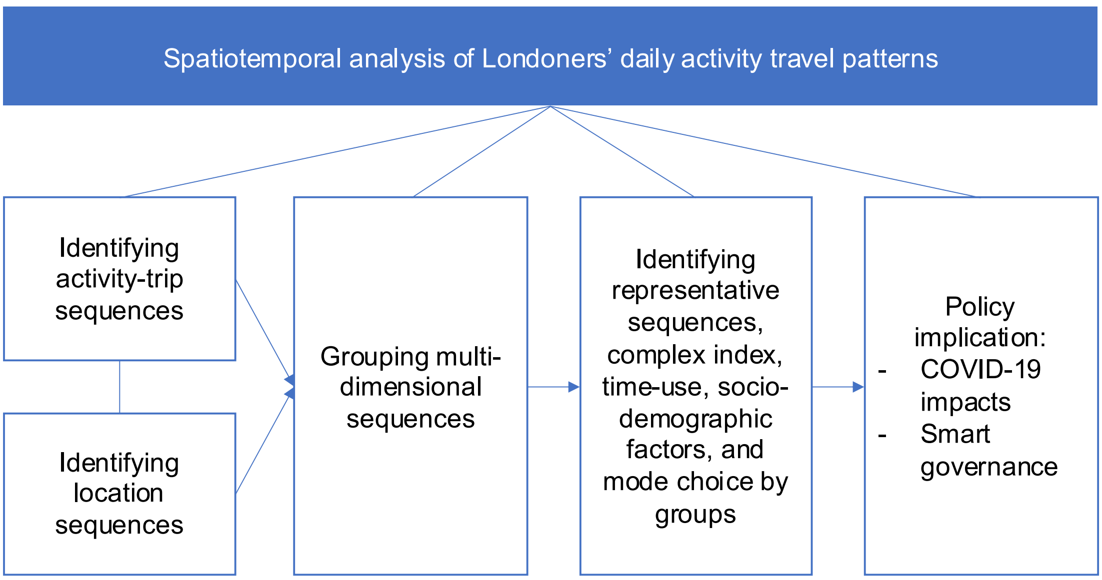

Figure 4-1: Analytical framework in Chapter 4

4.3.2 Study area and data

This chapter zooms into Greater London, the smartest city in transport development. The LTDS 2014/15 to 2019/20 are used as primary data for this analysis (TfL, 2020b). The LTDS is a continuous household survey in London area that collects information on household, individual, and all trips during designated travel days through three questionnaires (household questionnaires, individual questionnaires, and travel diaries). In the travel diaries, each member of the household reports all trips on the same day with detailed information, including trip purposes, modes, start and end times, locations of origins and destinations (TfL, 2011). The LTDS provides a representative sample of Londoners based on a random sampling of postcode addresses. The annual sample size of interviewed households is around 8000 (Fairnie et al., 2016). The data is robust enough to reveal dynamic patterns in London with three or more years of the sample (TfL, 2011). It is a reliable source and is widely used to analyse travel patterns and inform policymaking (e.g., the Mayor's Transport Strategy) in London (TfL, 2020e).

To further analyse the impact of COVID-19 from the LTDS 2019/20, we defined the period until the end of 2019 as the pre-pandemic period. Similar to the analysis in London report 13, the trips in 2020 are considered as travel during the pandemic (TfL, 2020e). Until the time of writing this thesis (March 2022), the latest data in the LTDS only covers the period up to March 2020. Thus, we only investigate the impact of COVID-19 at the early stage in this chapter.

GLA advised citizens to work from home since 17th March. Schools and pubs were closed on 21st March and the UK entered the first national lockdown on 24th March (GLA, 2020). Before the official measures, some Londoners were alerted about the potential outbreak of COVID-19 and the public awareness may lead to behavioural changes (Cheng et al., 2021; Li et al., 2021). Therefore, most of the daily sequences in 2020 in this study can be seen as self-organising patterns before strong restrictions. This chapter mainly analyses the early self-organising patterns after the pandemic.

4.3.3 Data pre-processing

4.3.3.1 Recoding activities, trips, and locations

An activity-travel sequence contains activities and trip episodes of an individual within a period. Following the activity categories in Allahviranloo and Aissaoui (2019), we grouped activities into eight types: work, education, shopping, personal activities, recreational socialising, home-based activities, other activities, and trips (Table 4-1). For trips, they were further divided into four types according to their main modes (based on distance).

Apart from the activities and trips, we also included location information to show spatial dynamics. We used land-use types of trips in the LTDS as the main types of location information. We reclassified the eleven types of land uses in the LTDS into eight groups: 1) residential, 2) office and factory/warehouse, 3) school/college, 4) shops, 5) public buildings, 6) open space, 7) health services (including hospital, GP/Dentist/other health services), and 8) other (including place of worship). We further added location information to the eight location types based on three geographies[4]: Inner London[5],

[4] We used the HIOX categories (01 Inner London, 02 Outer London, 03 External London) in the LTDS

[5] Includes 14 boroughs: Camden, City of London, Hackney, Hammersmith & Fulham, Haringey, Islington, Kensington & Chelsea, Lambeth, Lewisham, Newham, Southwark, Tower Hamlets, Wandsworth, and Westminster

Outer London[6], and Outside/External London[7]. This resulted in 24 location groups that contain both geographic and land use information. The detailed descriptions of each location type are listed in Table 4-2.

[6] Includes 19 boroughs: Barking & Dagenham, Barnet, Bexley, Brent, Bromley, Croydon, Ealing, Enfield, Greenwich, Harrow, Havering, Hillingdon, Hounslow, Kingston upon Thames, Merton, Redbridge, Richmond upon Thames, Sutton, Waltham Forest

[7] Includes commuter boroughs near Greater London (e.g., Dartford, Elmbridge, Epping Forest, Epsom & Ewell, Hertsmere) and the areas outside London

Table 4-1: Activity and trip types

Recoded activity	Recoded activity label	Description
Work	W	Work at usual workplace, delivery or loading work, other work
Education	E	Education
Shopping	S	Shopping food, other shopping
Personal activities	P	Personal business or use services, health or medical visit, participate in sport
Recreational socialising	R	Recreation or entertainment, stay at hotel or holiday home, worship or religious observance, leisure trip, visit friends or relatives at home, other social activities
Pick up drop off	K	Pick up or drop off someone to or from work, school, health visit, other place
In home	H	In home
Other	O	Other activities
Trip	T	Trip
Trip by private transport	TV	Trip by private transport
Trip by public transport	TB	Trip by public transport
Trip by cycling	TC	Trip by cycling
Trip by walking	TW	Trip by walking

Table 4-2: 24 location groups and codes

Administrative location	**Code**	**Description**
Inner London	IR	Inner London residential areas
	IW	Inner London office and factory/warehouse
	IC	Inner London school/college
	IS	Inner London shops
	IP	Inner London public buildings
	IG	Inner London open space
	IH	Inner London health services (including hospital, GP/Dentist/other health services)
	IO	Inner London other (including place of worship, and other)
Outer London	OR	Outer London residential areas
	OW	Outer London office and factory/warehouse
	OC	Outer London school/college
	OS	Outer London shops
	OP	Outer London public buildings
	OG	Outer London open space
	OH	Outer London health services (including hospital, GP/Dentist/other health services)
	OO	Outer London other (including place of worship, and other)
Outside Greater London	ER	Outside London residential areas
	EW	Outside London office and factory/warehouse
	EC	Outside London school/college
	ES	Outside London shops
	EP	Outside London public buildings
	EG	Outside London open space
	EH	Outside London health services (including hospital, GP/Dentist/other health services)
	EO	Outside London other (including place of worship, and other)

4.3.3.2 Creating daily sequences

The activity-travel information was then converted into daily travel sequences. In the LTDS, a travel day is defined as a 24-hour period that starts at 04:00 am and ends at 04:00 am on the following day. This chapter used the same definition of a travel day in the LTDS to construct daily activity-travel sequences. A daily sequence contains 288 intervals from 04:00 am to 04:00 am, in the form of a sequence of 288 letters. Each letter represents activity/trip or location state in a 5-minute interval.

In the LTDS, each respondent has several trip episodes in a day, containing trip ID, states, start times and end times. After data cleaning, an individual's trips on a travel day were converted into three types of daily travel sequences. The first type was used for preliminary analysis, containing seven activities and trips. The second type includes seven activities and four types of trips (based on the main mode). The last type contains spatial information. For spatial sequences, as it is hard to track the location changes every 5 minutes during the trips, we used the location type of destinations after the start of the trip.

For example, the daily sequences of an individual can be H/101-TV/2-R/28-TV/2-H/155 and OR/101-OP/30-OR/157 in the activity-trip channel and location channel respectively. The activity-trip sequence of H/101-TV/2-R/28-TV/2-H/155 means a person stays at home (H) for 101 intervals (i.e., 505 minutes since 4:00 am) and then spends 2 intervals (10 minutes) travelling by private transport (TV). He/she participates in recreational activities for 28 intervals (140 minutes) and goes back home by private transport in another 2 intervals (10 minutes). Then the person spends the rest of the day (155 intervals) at home until 4:00 am the next day. The location sequence of OR/101-OP/30-OR/157 shows that the individual stays in the Outer London home (OR) for 101 intervals and then goes to a public space in Outer London (OP). After spending 30 intervals (150 minutes) in the public space, he/she goes back to a residential place in Outer London (OR).

Altogether, the LTDS data we used in this chapter contains 204,108 trips and 68,716 sequences stemming from 68,716 individuals from 2015 to 2020. In the LTDS, each person has an interim expansion factor for the weekday sample. To explore the whole picture of activity-travel patterns in Greater London, we used the interim expansion factor to weight sequences.

To control the yearly difference, we carried out a preliminary clustering analysis comparing representative sequences of the different years. Preliminary results (see Appendix B-3) did not show significant differences in the years between 2015 to 2019. The main spatiotemporal trip-activity patterns in each year are similar. Thus, we analysed the daily sequences from 2015 to 2019 to understand the activity-travel patterns in London before the pandemic. We selected 50% of all sequences on weekdays as a sample set to find representative sequences. The year 2020 (first three months only) was analysed separately as we assumed some changes occurred after the pandemic.

According to the literature and preliminary results (in Appendix B-3), activity-travel patterns on weekdays and weekends were significantly different, so we split all sequences into weekdays' sequences and weekends' sequences (Millward et al., 2019; Su et al., 2020; Dharmowijoyo et al., 2017). From 2015 to 2019, we identified 48,142 daily sequences on weekdays and 18,062 sequences on weekends. In the first three months of 2020, 1910 daily sequences were identified on weekdays and 602 sequences were found on weekends.

4.3.4 Analysing multi-channel sequences

4.3.4.1 Calculating sequential complexity indices

Four indices were calculated to show the complexity of trip activity trajectories. The four complexity indices are transition, longitudinal entropy, turbulence, and complexity indices. The transition index computes the number of state changes in each sequence. The entropy index computes the longitudinal Shannon entropy of each sequence, reflecting the uncertainty in predicting states in a given sequence (Gabadinho et al., 2009). The turbulence index accounts for the number of unique sub-sequences in a trajectory and the difference of times spent in distinct states, which can illustrate within sequence turbulences (Elzinga and Liefbroer, 2007; Ritschard). Finally, the complexity index combines the number of transitions in the sequence and the longitudinal entropy (Ritschard; Gabadinho et al., 2011). It comprises both trip/activity states and the time period of activities in a daily sequence. The means of these indices were calculated to illustrate the complexity of daily sequences.

4.3.4.2 Calculating sequence dissimilarities

To align the sequences, the dissimilarities between each pair of sequences need to be first measured. The sequence alignment method is a method from molecular biology for analysing genetic sequences and is now one of the most widely used similarity measures in transport research (Shou and Di, 2018). It measures the distance between two sequences by calculating the minimal cost to transform one into the other (Kim, 2014; Wilson, 1998a).

Previous research has compared different distance calculation metrics and weighting schemes in measuring the dissimilarity between sequences (Song et al., 2021; Allahviranloo et al., 2017; Studer and Ritschard, 2016). The performances of different distance metrics are context dependent. Thus, a preliminary analysis was first conducted to decide the final metric.

Main methods for calculating distances between sequences include Hamming distance (HAM), time-sensitive Dynamic Hamming distance (DHD), Optimal matching (OM), Localised OM that emphasises adjacent states (Omloc), and OM between sequences of transitions (Omstrans) (Song et al., 2021; Allahviranloo et al., 2017; Studer and Ritschard, 2016). Studer and Ritschard (2016) examined the sensitivity of different metrics to changes in states (i.e., activity), timing (i.e., the position of the states in the sequence), duration (i.e., consecutive times spent in a state), and sequencing (i.e., the order of activities). The results showed that Hamming-based distances (HAM and DHD) are sensitive to timing. Metrics from the Omloc family are more sensitive to duration and sequencing. OM distances are less sensitive to the difference in temporality than sequential orders. Song et al. (2021) found the results of HAM and OM are similar while DHD is more sensitive to temporal changes (Song et al., 2021).

Most previous studies have applied unit cost weighting in measuring dissimilarities among sequences while recent studies have suggested weighted distance. For example, Song et al. (2021) compared unit-cost, fixed-flexible weighted, trip-activity weighted, and transition-based weighted distance through clustering results. They found that the choice of weighting schemes can have a greater influence than the choice of distance metrics on clustering results. Three main methods to choose substitution costs are theory-based, state attribute-based, and data-driven costs (Studer and Ritschard, 2016). Data-derived cost OM is more sensitive to changes in duration, small changes in rate events and perturbations. A popular method of generating data-driven weight is to derive transition rates from observed cases (Studer and Ritschard, 2016). Transition cost derived from empirical data can better illustrate patterns on changing types of activity while time-dependent transition-based weighted is more sensitive to temporal changes (Song et al., 2021).

In the preliminary analysis (see Appendix B-1), we calculated pairwise distances using four metrics with different weighting schemes below. The four distances we used are 1) unit cost OM, 2) transition-based OM, 3) localised OM that accentuates adjacent states, and 4) DHD that considers time-dependent transition rate between two states. The pairwise distances were calculated through R package *TraMineR* (Gabadinho et al., 2011). Four distance calculating methods were applied to the simplest 9-state activity-trip sequences. We selected 10% of basic activity-trip weekday sequences from 2015 to 2019. In total, 4814 sequences were assessed, using four distance matrices and the most used Ward clustering method. Clustering results and computation times were compared. Based on the preliminary result in Appendix B-1, we selected transition-based OM distance for further analysis in this chapter.

To analyse both activity-trip and location information of each person in each time interval, we computed multichannel pairwise distances. The multichannel approach can assess multidimensional information simultaneously. The multichannel dissimilarity matrix sums the transformation costs of each channel in the multistate sequence (Gabadinho et al., 2011; Eisenberg-Guyot et al., 2020). In this case study, both activity and location of a person in each interval were considered.

4.3.4.3 Clustering daily sequence

To find groups from the pre-defined dissimilarly matrix, this study tested two types of clustering techniques in the preliminary study (see Appendix B-2). The first technique is agglomerative clustering with Ward linkage (i.e., Ward method), which is a widely used hierarchical clustering approach in transport studies (Song et al., 2021; Cho et al., 2019). The Ward's method is a bottom-up clustering approach that minimises the variations within clusters (Ward, 1963). The result produces a tree structure named dendrogram to represent the hierarchy of clustering. Ward clustering does not require a predefined number of clusters and tends to generate clusters of similar size (Ezugwu et al., 2021). However, hierarchical methods are less sensitive to outliers (Xu and

Wunsch, 2010).

Fuzzy clustering is based on the concepts and ideas from fuzzy-set theory (Ruspini et al., 2019). As fuzzy-set theory does not believe in definite belongingness, fuzzy clustering discovers the degree of each data point to several clusters (Winkler et al., 2010). Fuzzy clustering has been less applied in sequence analysis, mainly because of technical barriers (Studer, 2018). Relaxed assumptions of fuzzy clusters that allow a sequence to belong to more than one cluster can provide more interesting and adaptive results. Activity-travel trajectories can be highly uncertain and complex so they can be categorised into more than one type. This study used three fuzzy clustering methods of daily sequences in the preliminary analysis.

A pre-defined dissimilarity matrix can be accepted in FANNY and NEFRC algorithms. The three fuzzy methods we tested were FANNY clustering, NEFRC clustering and NEFRC with noise. FANNY algorithm was proposed by Kaufman and Rousseeuw (2009). We used R library *Cluster* for FANNY clustering (Maechler et al., 2013). The other fuzzy methods were conducted with R package *Fclust* (Ferraro et al., 2019). The NEFRC is a non-Euclidean fuzzy relational clustering algorithm that can additionally consider noise clusters that contain all outliers (Dave, 1991). The detailed equations of the fuzzy clustering algorithms can be found in Ferraro et al. (2019).

In the preliminary analysis in Appendix B-2, four clustering methods were applied to the simplest 9-state activity-trip sequences. 5% of activity-trip weekday sequences (2405 sequences) from 2015 to 2019 were selected to compare clustering methods. Silhouette index that measures the compactness of clusters is used to assess clustering results (Ezugwu et al., 2021). We conducted an automatic comparison of Ward and FANNY clustering results with different cluster numbers using the R package *ClValid* (Brock et al., 2008). Silhouette index of NEFRC and NEFRC-noise clustering results were extracted through R package *Fclust* (Ferraro et al., 2019). Based on the preliminary result of clustering comparison, we selected FANNY clustering for further analysis in this chapter.

4.3.4.4 Finding representative patterns

We used two methods to find representative patterns of clustering results. We first extracted the sequence medoid of each cluster, using the R package *WeightedCluster* (Studer, 2013). The medoids are representative sequences within clusters and have the minimum sums of distances between representative sequences with other sequences.

Additionally, we plotted the graphical representation of fuzzy clustering results in each cluster based on the fuzzy membership degree (Studer, 2018). A threshold of 60% was chosen when plotting the sequences. It means the sequences with a minimum possibility of 60% can be included in the final plots of fuzzy clustering results. The plots were generated through the R package *WeightedCluster* (Studer, 2013).

4.3.5 Finding socio-demographic determinants

Socio-demographic variables such as age, gender, household structure and income are important determinants of travel patterns, as shown in section 4.2.2. In this case study, we investigated the variables extracted from the LTDS (see Table 4-3). It should be noted that the working from home status is not included in the basic socio-demographic models in exploring cluster profiles before 2020.

We first calculated the frequencies in each variable. Multicollinearity was checked through calculating variance inflation factors. After dealing with multicollinearity issues, we built Dirichlet regression models.

For fuzzy clustering results, the cluster membership is represented by a membership matrix. Under this circumstance, the widely used multinomial regression models cannot be applied to analyse explanatory variables. As suggested by Studer (2018), Dirichlet regression models can be built to examine the impacts of key factors on clustering results. Similar to the widely used multinomial models, one cluster is chosen as the reference and the Dirichlet regression model analyses the influence of different variables on the likelihood to be in this cluster rather than in other clusters (Studer, 2018). We chose the largest cluster as the reference group in this study. The coefficients of covariates in the model can show the effect of explanatory factors on clustering. Dirichlet Regression models were run in R package *DirichletReg* (Maier, 2014).

Table 4-3: Selected socio-demographic variables

Categories	Attributes	Codes	Subcategories description
Household	Accessible vehicles	C0	0
		C1	1
		C2	2 or more
	Household structure	H1	Couple with children
		H2	Couple without children
		H3	Lone parent
		H4	Single adult
		H5	Single pensioner
		H6	Other
	Household income	IL	Low-income household
		IM	Middle-income household
		IH	High-income household
Person	Gender	M	Male
		F	Female
	Age	A1	5-15
		A2	16-24
		A3	25-44
		A4	45-59
		A5	60+
	Ethnic group	E1	White
		E2	Mixed, Other and Arab
		E3	Asian
		E4	Black
	Driving license holder	D	Yes
		ND	No
	Oyster card holder	B	Yes
		NB	No
	Occupation	O1	Full-time worker
		O2	Part-time worker
		O3	Student, school pupil and teenager

		O4	Unemployment
		O5	Unable to work because of long-term illness or disability
		O6	Retired
		O7	Looking after home
		O8	Other
	Health condition	H	No health problem
		HP	Long term health problem that limits activity
		HM	Mental health condition
	Working from home status	WFH	Usually working from home
		NWFH	Do not working from home

4.3.6 Assessing the COVID-19 impacts

Our dataset contains the travel activities from January to March 2020, which we used to analyse the early impact of COVID-19 on trips and activities in London. We defined daily sequences between 2015 to the end of 2019 as the pre-COVID sequences and sequences in 2020 as COVID-influenced sequences. As the data in 2020 only contains sequences from January to March, we created a subset of the pre-COVID group by filtering sequences in January, February, and March to avoid seasonal impacts on travel behaviours. Thus, the pre-COVID group contains sequences from January to March from 2015 to 2019.

We first calculated the complexity indices and the time used in each activity and trip in different main modes to show the difference before and during the early stage of the pandemic. We then clustered two sets of data. Representative sequences were then compared. Working from home status was added to the socio-demographic variables in the baseline models to analyse the role of working from home status before and during the pandemic. Analysis of variance (ANOVA) tests were run to analyse the impacts of the COVID-19 related variables by comparing regression models with and without the COVID-19 related variable.

4.4 Results

Eleven clusters on weekdays and eight groups on weekends are found through fuzzy clustering. The socio-demographic characteristics in each cluster are presented in section 4.4.1. We named each cluster based on the typical activity patterns and the main socio-demographic features. The impact of the COVID-19 pandemic on individual trip-activity patterns is shown in section 4.4.2.

4.4.1 Activity-travel patterns in recent years

4.4.1.1 Representative sequences on weekdays

The weekdays' sequences can be grouped into eleven fuzzy clusters. The main activities are work, home-based activities, education, and recreational activities. Public transport is the main mode of the day work and day education clusters. The stay-at-home clusters contain short trips and activities, mainly recreation and shopping at midday. The most important socio-demographic determinants on weekdays' clusters are age, occupation, household income and car access.

Based on the modelling results and the proportion of each variable in different clusters, the main socio-demographic profiles of each cluster are built. The details of the model output and frequency table are in Appendix C (Table C-1 & Table C-2). The two student clusters, five worker clusters and four non-worker/student clusters are as follow.

Cluster #1: Outer London youths, involved a group of students who participate in education during the day and go back homes in Outer London. 9.4% of daily sequences are in this group. Most students go to schools by public transport and walking. The representative patterns are H/52-TW/1-E/84-TW/1-H/150 and OR/52-OC/85-OR/151. A typical youth leaves the Outer London homes at around 8:20 am and walks to school for 5 minutes. He/she spends most of the day at school (around 7 hours) and goes back home at 5:25 pm.

Most youths are younger than 15 years old, which count for 81% of people in this cluster. A smaller proportion (16%) of students are between 16 to 24 years old. Black, Asian and minority ethnic (BAME) Londoners count for 47% of members. 21% of them are Asian, 14% of them are black Londoners, and 12% of them are from other minority groups. Almost all of them (97%) are in a healthy status. Around half of their households are middle-income families with at least one accessible vehicle. Most of their households are labelled as "couples with children".

Cluster #2: Inner London youths, involved a group of students who go to schools or colleges during the day by public transport and walking. The size of this cluster (6.4%) is smaller than that of Cluster #1. The representative sequences in this cluster are H/53-TW/1-E/83-TW/1-H/150 and IR/53-IC/84-IR/151. A representative young Londoner in this group goes to a school in Inner London at 8:25 am on foot. After spending around 7 hours studying at school, the young student walks back home in 5 minutes and stays at his/her Inner London home for the rest of day.

Similar to Cluster #1, most students are below 15 years old, counting 73% of people in this cluster. 19% of members are between 16 to 24 years old, which is more than that of Cluster #1. In terms of ethnicity, Inner London youths are more ethnically diverse than the Outer London youths. The BAME Londoners are likely to be in this cluster, with 52% of members. 22% of them are black, 18% of them are Asian, and 12% of them are from other minority ethnic groups. 96% of them are in a healthy status.

Compared to Cluster #1, more youths are from lower-income households. People in this cluster are generally from middle-low-income families, of which 35% of them are from low-income households. 47% of their household have one accessible vehicle while 44% of them do not have access to any vehicle. Similar to Cluster #1, a great proportion (73%) of the households are classified as "couples with children".

Cluster #3: Mixed place day workers living in Outer London, consisted of a group of full-time and part-time workers who engage in working activity during the day. Their main working places are schools/colleges and public buildings in Outer London, as well as offices outside London. The main modes in this cluster are private and public transport. This is the largest cluster among workers' clusters, counting for 11.7%. The typical sequences are H/45-TV/2-W/115-TV/2-H/124 and OR/288. The activity sequence means a typical individual in this group leaves home at 7:45 am and then spends 10 minutes travelling to work by private transport. He/she works for 9 and a half hours and at around 5:00 pm goes back home by private transport (10 minutes), staying at home for the rest of the day. The long hour "work" activity contains the time for rest breaks during the working hours. As the working places are complex in this group and some of them work from home, the representative location sequence in this cluster shows that people are mainly at home for the whole day. This might explain the long time (over 9 hours) spent on "work" activity.

About half of the people in this cluster are between 25 to 44 years old and another one thirds of people are between 45 to 59. Most members are full-time workers while another 12% of them are part-time workers. Almost all of them (98%) are healthy. 75% of the individuals in this cluster hold a driving license and 70% have an oyster card.

Most of them are from middle-high-income households. About half of them are from middle-income families and 41% of them are from high-income households. 38% of them are from a household with two or more accessible vehicles. Their main household structures tend to be classified as "couples without children" (37%) or "single pensioners" (14%).

Cluster #4: Mixed place day workers living in Inner London, involved a group of full-time workers who engage in working activity during the day. This is the second-largest cluster among workers' clusters, with 9.3% of people. Outer London offices, Inner London public buildings, and Inner London offices are the main working places in this cluster. Public transport is the most used commuting mode. H/50-TB/6-W/105-TB/6-H/121 and IR/50-OW/111-IR/127 are the representative sequences. A typical worker in this group leaves his/her Inner London home at 8:10 am and goes to work by public transport, spending half an hour. He/she spends around 8 hours and 45 minutes at Outer London offices, mainly working. At around 5:25 pm, he/she goes back home in Inner London by public transport (30 minutes) and stays at home for the rest of the day.

71% of them are white and 58% of them are between 25 to 44 years old. The people in this cluster are younger and less ethnically diverse than workers in Cluster #3. 70% of them hold an oyster card and two thirds of the people hold a driving license. Almost all of them are in healthy condition.

Similar to those in Cluster #3, they are generally from middle-high income households. 44% of them are from a high-income household and 47% of them belong to the middle-income range. Half of them do not have any accessible vehicle. 36% of their household structure is a couple without children. A smaller number of them are lone parents.

Cluster #5: Outer London day workers, involved a group of full-time workers who work in Outer London offices or factories and live in Outer London residential areas. This cluster counts for 5.5% and the main mode is public transport. The representative sequences are H/47-TB/6-W/108-TB/6-H/121 and OR/47-OW/114-OR/127. The activity sequence means a typical individual in this group leaves home at 7:55 am and then spends 30 minutes travelling to work by public transport. He/she works for 9 hours and at around 5:25 pm goes back home by public transport (30 minutes), staying at home for the rest of the day. It should be noted that the 9-hour work activities contain the time for short breaks at work, including resting at offices, very short shopping, and recreation activities at midday. The location sequence shows that a typical person lives in Outer London residential areas and goes to work in Outer London offices during a working day.

Males are more likely to be in this group, counting for 56%. The ethnic groups are more diverse than other workers' groups, with 22% Asian Londoners. Nevertheless, 64% of them are white. Half of the people are between 25 to 44 years old and one-third of them are between 45 to 59 years old. Compared to other workers' groups, more middle-aged workers are found in this group. Almost all of them are in healthy condition. A large proportion of the individuals in this cluster hold a driving license and 71% hold an oyster card.

More than half of their families are high-income households. 36% of their household can access more than two vehicles. Their household structure is more likely to be a couple without children.

Cluster #6: Inner London day workers, comprised of a group of full-time workers who engaged in working activity during the day. Most of them work in Inner London offices or factories and live in Inner London residential areas. The size of this cluster (6.6%) is slightly larger than Cluster #5. The main mode in the cluster is also public transport. H/50-TB/6-W/111-TB/6-H/115 and IR/50-IW/117-IR/121 are representative sequences in two channels. A typical worker in this cluster leaves his/her Inner London home at 8:10 am and goes to work by public transport (30 minutes). He/she spends around 9 hours and 15 minutes at the Inner London office, mainly working. The worker then goes back home in Inner London at 5:55 pm by public transport (30 minutes) and stays at home for the rest of the day.

The members in this group are mainly white and wealthy, with 77% white Londoners. This cluster is the least ethnic diverse among all workers' groups. 54% of them are male. Most people (68%) in this cluster are between 25 to 44 years old. Almost all of them are in healthy condition. 72% of the individuals in this cluster hold a driving license and 66% of them hold an oyster card.

The annual income is highest among all groups, with 60% of them from high-income households. Nevertheless, more than half of their households do not own any vehicle. 40% of their household structure is “couple without children” and a smaller percentage is “lone parent”.

Cluster #7: Inner London day workers living in Outer London, involved a group of full-time workers who work in Inner London offices or factories and live in Outer London residential area, mainly commuting by public transport. This is the smallest cluster (4.8%) among all workers’ clusters. Compared to other workers’ clusters, people in this cluster have longer commuting time. H/45-TB/10-W/106-TB/12-H/115 and OR/45-IW/116-OR/127 are the representative activity and location sequences respectively. An individual in this cluster typically leaves his/her Outer London home at 7:45 am and goes to work by public transport, spending 50 minutes commuting. He/she spends nearly 9 hours at the Inner London office and goes back home at 5:25 pm by public transport (50 minutes), staying at the Outer London home for the rest of the day.

This cluster has the largest proportion of males among all groups. 57% of them are male. 62% of cluster members are between 25 to 44 years old. Two thirds of them are white and 20% of them are from Asian backgrounds. Almost all of them are healthy. 83% of the individuals in this cluster hold a driving license and 76% hold an oyster card.

59% of their families are high-income, which is the second wealthiest among all groups. Around half of people in this cluster have one accessible car in their households. 39% of members in this cluster are from couple families with children while another 38% of them are from couples without children’s households.

Cluster #8: Outer London stay-at-homes, consisted of a group of people who spend most of their time at their Outer London homes. Most of them go out at midday for less than an hour. The main out-of-home activity is shopping at Outer London shops at midday, mainly using private transport or walking. This cluster has the largest membership (16.6%) among all clusters. The representative sequences in this cluster are H/288 and OR/288, meaning the people mainly participated in home-based activities at their Outer London homes.

56% of members in this cluster are female and 38% of cluster members are the elderly (above 60 years old). 68% of them are white people - more than that in Cluster #9. The health status is worse, with 13% of people suffering from long-term health problems that limit their activities. The main working status of members are retired (32%), looking after home (12%), and part-time working (11%).

Most people come from middle-low-income households. 48% of people in this group come from middle-income households and 26% of them come from lower-income households. About half of them are from a household with one accessible vehicle. Their household structures tend to be “couples without children” or “single adults”.

Cluster #9: Inner London stay-at-homes, involved a group of people who spend most of their time at home during the weekdays. Most of them go out for very short shopping or recreation in shops and public buildings within Inner London. The main modes are public transport and walking. It has the second largest membership, with 15.6% sequences in this cluster. The representative sequences are H/288 and IR/288. In this cluster people mainly stay in their Inner London homes, conducting home-based activities.

This cluster has the largest proportion of females among all groups, with 57% female. 29% of the cluster members are above 60 years old. The ethnic backgrounds are more heterogeneous, with 16% of members being black. A large proportion of people in this cluster are non-workers or part-time workers. Retired people take up 23% and people who mainly look after the home count for 11%. The health status is worse in this group, with 12% of people suffering from long-term health problems.

As for household characteristics, more than half of people cannot access any vehicle and 31% of cluster members are from low-income households. Compared to Cluster #8, the people in this group are generally from lower income families. Lone parents, single pensioners, and single adults are more likely to be in this cluster.

Cluster #10: Outer London afternoon recreation, comprised of a group of part-time workers and non-workers who conducted entertainment activities or social activities in the afternoon, starting from 1:00 pm and ending at around 3:00 pm, mainly using public transport. H/101-TV/2-R/28-TV/2-H/155 and OR/101-OP/30-OR/157 are the representative sequences showing that individuals tend to entertain in public places in Outer London for around two and a half hours in the afternoon.

54% of people in this cluster are women and one third are above 60 years old. 69% of them are from white ethnicity. 64% of the individuals in this cluster hold a driving license while more than half are not oyster cardholders. The main working status of them are retired (26%) or part-time workers (14%). Compared to other clusters, the health condition is also worse in this group, with 11% of people being physically unwell. Most of them have middle range annual income with two or more accessible vehicles. In terms of household structure, 37% of them are couples without children. 12% of them are single pensioners and 8% of them are single adults.

Cluster #11: External London mixed activities, consisted of a group of people who spend most of their time at their homes outside Greater London. The mixed activities include day work and day education. Public transport is the main mode of travelling. This cluster has the smallest membership (3.2%). The representative sequences in this cluster are H/288 and ER/288.

A large proportion of members in this cluster are white and 54% of members are female. The main occupations of members are full-time workers (46%), part-time workers (10%), students (21%), and retired people (12%). The health status is worse in this group, with 7% of people suffering from long term physical health problems.

Half of the people are from middle-income households. 29% of them can access at least two vehicles. Their household structures vary.

4.4.1.2 Representative sequences on weekends

We identified eight clusters on weekends. The weekends' sequences are less complex than those of weekdays. The main activities are home-based activities, recreational activities, and work. The determinants of weekends' patterns are household structure, household income, accessible vehicles, and ethnic group, which are slightly different from those of weekdays. Based on the data mining results (see Tables C-3 & C-4 in Appendix C), the main characteristics of each cluster are as follows.

Cluster #1: Outer London stay-at-homes, involved a group of people who spend most of their time at home during the weekends. Most of them go out for very short shopping or recreation, mainly in Outer London shops. This is the largest cluster at weekends, with 24.3% sequences. Private transport is the main mode in this cluster. The representative sequences are H/288 and OR/288, meaning people basically stay at their Outer London homes.

A large proportion of members in this group are in middle and old ages. Particularly, 22% of them are between 45 to 59 years old and 24% of them are above 60 years old. Compared to Cluster #2 and #3, more elderly can be found in this group. People who are retired count for 19% of this group. 67% of them are white and 19% are from Asian backgrounds. The health status is worse in this group, with 8% of people having long term physical health problems.

About half of them are from middle-income families with at least one accessible car. In terms of household structure, 42% are labelled as “couple with children” and 34% are “couple without children”.

Cluster #2: Inner London stay-at-homes, involved a group of people who spend most of their time in their Inner London homes. The main out-of-home activities are recreation and shopping at Inner London shops, using walking as the main mode. This is the second-largest cluster on weekends, with 18% of sequences. The representative patterns are H/288 and IR/288.

Compared to Cluster #1, people are from more diverse backgrounds, with 16% being black Londoners. The health status in this group is as worse as in the first cluster. People who mainly look after homes are more likely to be in this cluster. People are from poorer households. 21% of people come from low-income households and about half of them cannot access any vehicle in their household. Their household structures are more likely to be lone parents or single adults.

Cluster #3: External London stay-at-homes, comprised a group of people who spend most of their time at homes outside London. Most of them go out for short recreation or shopping, mainly in External London shops. Their main mode of travelling is private transport. H/107-TV/1-K/1-TV/1-H/178 and ER/107-EO/2-ER/179 are typical sequences in activity and location.

Most people are white in this group and 54% of members are female. About half of them are full-time workers and 65% of them are holders of a driving license. Compared to Cluster #1 and #2, people are from higher income households. Most people come from middle-high-income households, with at least one available car. 38% of their household structure are couples without children.

Cluster #4: Outer London afternoon-night recreation, involved a group of people who participated in recreational activity mostly in Outer London homes or public buildings, from 3:00 pm to 7:00 pm. The main mode in this cluster is private transport. H/131-TV/1-R/53-TV/1-H/102 and OR/288 are representative activity and location sequences respectively.

About one fifth of the members in this group are students. 57% of them are oyster cardholders. Asian Londoners tend to be in this cluster. The household annual income of about half of people is in the middle range. 34% of cluster members have at least two available vehicles in their households. 41% of their household characteristic is the couple with children.

Cluster #5: Outer London afternoon recreation, consisted of a group of people who participated in recreational activity mostly in Outer London public buildings or open spaces, from 12:00 pm to 3:00 pm. They mainly use private transport for their trips. The sequential exemplars are H/95-TV/2-R/40-TV/2-H/149 and OR/288.

About half of the people in this cluster are female. 20% of cluster members are below 15 years old. Students or retired people are more likely to be in this cluster. Most of them can access at least one vehicle. 46% of them are accessible to one vehicle and one third of people have two or more available cars in their households. Half of the members are from a family that can be characterised as a “couple with children”.

Cluster #6: Inner London afternoon-night recreation, consisted of a group of people who participated in recreational activities, starting from 2:00 pm to 7:00 pm. Their main recreational places are Inner London homes, Inner London public buildings, Outer London public buildings, and Inner London open spaces. The main mode is public transport. H/119-TW/1-R/59-TW/1-H/108 and IR/288 are the exemplary sequences in the two channels.

48% of people in this cluster are between 25 to 44 years old. 69% of them are white Londoners. 58% have an oyster card. Most people in this cluster are full-time workers, counting 53%. 41% of members are from high-income families. But half of them cannot use any vehicle in their households. For household structure, lone parents are more likely to be in this cluster.

Cluster #7: Outer London day workers, comprised a group of workers who engaged in work activity in Outer London from 9:00 am to 4:00 pm during the weekends. Their main working places are Outer London offices/factories, Outer London shops and Outer London public buildings. The main mode in this cluster is public transport. The representative activity-trip pattern is H/53-TV/1-W/89-TV/1-H/144 and the typical location sequence is OR/53-OP/90-OR/145.

55% of the people in this cluster are males. The majority of people are between 16 to 59 years old. About two-thirds of them are full-time workers. The ethnic background is more diverse than Cluster #8, with19% Asian Londoners. Almost all of them are in healthy status. 64% of them have an oyster card and 64% of them hold a driving license. The household annual income of the members in this cluster is mainly in the middle range. 35% of them can access two or more vehicles in the household. Most of the household structure of members in this cluster are “couples without children” and a smaller number are “single pensioners”.

Cluster #8: Inner London day workers, involved a group of workers who engaged in short-hour work activity from 11:00 am to 3:00 pm during the weekends, mainly in Inner London public buildings, shops, and offices. The main mode in this cluster is public transport. The sequential exemplars are H/83-TW/2-W/52-TW/2-H/149 and IR/83-IP/54-IR/151.

The majority of people in this group are white. This group has the best health status among all clusters, with 97% in good health status. Most people are full-time workers between 16 to 44 years old. 62% of them are oyster cardholders. But about half of the people in this cluster do not have access to any vehicle in their households. The income level is middle-high, with 40% in high-income households. 19% of the household structure of members are lone parents, and a small number are single adults.

4.4.2 The COVID-19 pandemic impacts

Compared to previous years, more respondents of the LTDS stated that they travelled less in early 2020 (January to March 2020). The COVID-19 pandemic may have an early impact on individuals' decision of making fewer or no trips before the lockdown restrictions on the 24th of March 2020.

4.4.2.1 Complexity of daily sequences

As illustrated in Table 4-4, the turbulence index remains stable before and after 2020, indicating all years witnessed similar unique sub-sequences and time use in each activity. The transition index illustrates that the number of activity changes in 2020 is less than that of the previous years, both on weekdays and weekends. The entropy index shows that longitudinal entropy in 2020 is less, which means it is easier to predict sequences in 2020. The uncertainty level of the sequences has been reduced. Overall, the complexity level of the individual sequences is slightly lower than those of previous years. This may be explained by Londoners' self-organising behaviours towards the pandemic. Facing the emergent pandemic, people reduced unnecessary trips/activities and their usage of public transport, to avoid being infected in early 2020. Still, the complexity of individuals' daily sequences varies on weekdays and weekends.

Table 4-4: Complexity indices of 5-year trip-activity sequences

Days	Year	Number	Transitions	Entropy	Turbulence	Complexity
weekday	2015-2019	9614	5.770	0.319	0.026	0.077
weekend	2015-2019	3764	5.672	0.262	0.025	0.069
weekday	2020	1910	5.758	0.312	0.026	0.076
weekend	2020	602	5.640	0.256	0.025	0.068

4.4.2.2 Time-use in activities and trips

We calculated the mean time in each subset of the sequential data. Compared to the previous five years, people spent more time at home in early 2020. During the weekdays, the mean duration at home is 1.6% higher than that in the previous years. People also spent slightly more time shopping after 2020. Additionally, pick up and drop off activities took a longer time in early 2020 than in previous years. This may be because parents have an increased willingness to pick up and drop off their children. On the contrary, less time has been used in education, work, recreation, and personal business in the first three months of 2020. Early 2020 also witnessed less time spent on trips. Focusing on the four main modes, we found that the mean times of trips by private transport and cycling increased while that of public transport decreased. During the weekends, less time was spent on education, recreation, and pick up/drop off activities and trips. Londoners spent more time at home on both weekdays and weekends.

4.4.2.3 Representative sequences

To analyse the changes on weekdays, both sequences (in January to March) before and after 2020 were first classified into eleven clusters. The clustering results show that typical clusters before and after 2020 are mostly similar. Both include two stay-at-homes groups, a mixed activity group, two students' groups, an afternoon recreation group, and five workers' groups. It should be noted that the workers' groups of January to March sequences are slightly different from those of all months.

In early 2020, more sequences are classified into either the Inner London stay-at-homes or Outer London stay-at-homes. Besides, the External London mixed activities group that contains a large proportion of home-based activity also grew larger. As non-essential trips and activities reduced after 2020, the afternoon recreation cluster decreased. In terms of students' groups, fewer Inner London students went to schools/colleges while the percentage of educational activities in Outer London did not decline. This may be because the Inner London students contain a larger proportion of college/university students that are more flexible in adjusting their trips to universities. Younger students in primary and middle schools still needed to study at schools before the implementation of the government restriction to close schools on 21st March 2020.

The proportions of almost all workers' groups decreased in 2020, except for the mixed places workers in Inner London that contain homeworkers. The general decrease in out-of-home working activities may be related to governmental advice on work from home since 17th March 2020. The typical five workers' groups in January to March 2020 are: 1) mixed place workers living in Outer London, 2) mixed place shorter-hour workers living in Outer London, 3) mixed places day workers living in Inner London, 4) Outer London day workers, and 5) Inner London day workers living in Outer London. We found more people working in mixed places, which include homes.

The main modes of some groups changed. Before 2020, the primary mode for almost all groups is public transport. In early 2020, people tended to prefer private transport. The main mode of the mixed places workers living in Outer London and mixed activities outside London became private transport.

For the clusters on the weekends, we identified eight groups in the first three months of 2020. The representative spatiotemporal patterns and main modes changed. The eight groups in 2020 are: 1) Outer London stay-at-homes, 2) Inner London stay-at-homes, 3) Outer London afternoon-night recreation, 4) Outer London afternoon recreation, 5) Inner London midday recreation and afternoon shopping, 6) Outer London day workers, 7) Inner London afternoon-night recreation, and 8) External London stay-at-homes.

Within the eight groups, three stay-at-home clusters are similar to those stay-at-home groups before 2020. In the early stage, we only found one workers' group at the weekends, indicating the working activity was reduced on the weekends. The typical characteristics of this working group are 10:00 am to 5:00 pm day work in mixed places, including Outer London public buildings, offices, and Inner London offices.

The rest four groups are related to recreational activities, in the afternoon and night. A group of people participated in entertaining activities in the midday from 11:00 am to 2:00 pm and then shopping for around an hour (3:00 pm to 4:00 pm) before returning home. This may be because people would like to shop for more food and necessary commodities when facing the pandemic.

The main mode of each group also changed after 2020. From 2015 to 2019, the primary mode for the workers' group and the Inner London afternoon-night recreation group is public transport. In early 2020, the workers' group shifted their main mode of travelling to private transport. The Outer London stay-at-homes and Inner London stay-at-homes both preferred walking during their shop trips in 2020. The Inner London recreation groups still used public transport as the main approach to travel. Besides, private transport became the main mode of travelling for all the other groups.

4.4.2.4 Socio-demographic determinants in 2020

During weekdays in early 2020, the main personal determinants are ethnic groups and occupations. The most important household factor for weekdays' patterns is the number of accessible vehicles in the household. Ethnic groups, particular the Asian group, has a significant impact on individuals' activity-travel choices in early 2020. These determinants are different from those before the pandemic, which were age, occupation, household income and accessible vehicles. During weekends in 2020, the main personal determinants are age, ethnic status, oyster card holders' status, and mental health status. The most important household factors are the accessible vehicles and income levels in the household. These determinants are also slightly different from those before the pandemic (household structure, household income, accessible vehicles, and ethnic group).

For the COVID-19 related covariate (i.e., work from home status), this factor can significantly impact the choice of daily activity patterns on weekdays in early 2020. However, it is not a determinant of weekends' patterns after 2020. Furthermore, we tested this variable on pre-COVID datasets and found that it was also a significant determinant of weekdays' clusters. This indicates that the status of working from home has long been a determinant of travel behaviours and the pandemic has amplified its impact.

4.5 Discussion

We have presented a novel interdisciplinary approach for adaptive understanding daily activity-travel patterns, which can address the methodological gaps of lacking sophisticated approaches to understanding complexity in dynamic transport systems (Gap 2) and lacking the use of citizen-centric data and correct understanding of citizens' needs (Gap 3). It allows us to learn the travel behaviours of different population groups and transport market segmentation with acceptance of uncertainties. Here, our theoretical and practical research contributions, the added value of the proposed method, the case-specific insights and potential policy implications as well as the limitations and further directions are discussed.

4.5.1 A fuzzy understanding to support adaptive planning

Uncertainty has long been a concern in understanding, planning, and governing the transport sector. As illustrated in previous chapters, adaptive planning that considers various scenarios can help manage the uncertainties in urban transport. The fuzzy approach in this study can be an important tool for adaptive planning in managing the dynamic complexity of the changing activity-travel patterns. An important theoretical implication of this chapter is to discover the adaptive understanding of travel behaviours through fuzzy clustering of daily activity-travel sequences. In fuzzy clustering, membership is gradual, indicating a sequence has a certain possibility in different groups. From a complexity viewpoint, travellers may not be strictly rational customers with certain decisions. Understanding travel patterns through a fuzzy approach that recognises a level of uncertainty in travel behaviours can reveal the complexity of travel activities.

When facing the emergence of COVID-19, the fuzzy approach can better capture the adaptation and uncertainty in changing behaviours of travellers. For example, in the case study (London), we investigated the self-organising behaviours of the citizens before the strict mobility-restricted measures (e.g., lockdown). The minor changes in the early stage we extracted can be used to support adaptive planning in both crisis management and transport governance.

4.5.2 A robust approach for unfolding complexity

A methodological implication of this chapter is to provide a robust understanding of activity-travel patterns through an interdisciplinary approach. This chapter identified daily spatiotemporal activity-travel patterns with a novel interdisciplinary approach. Specifically, we used a fuzzy multi-channel sequence analysis that measures dissimilarity with transition-based optimal matching distances derived from travel survey data and clustered dissimilarity matrix through a fuzzy unsupervised learning algorithm. This chapter further linked the clustering results (i.e., output probability vectors) with important variables through descriptive analysis and Dirichlet regression models.

The method can provide a robust understanding of the complexity of the urban transport system and support smart transport governance through new data-driven evidence. The approach has been applied to the case of Greater London and generated interesting results of different groups of Londoners, capturing the flexible information and emergent changes due to COVID-19.

The method can be transferred to other cases and fields. The fuzzy spatiotemporal pattern recognition methods can be further applied to other transport data such as vehicle records and GPS records. This approach can be used to construct daily sequences of a person or vehicle, incorporating information on different aspects (e.g., location, activity, status, and emission). In addition to transport studies, the methodology can also be applied to understand other sequential data such as transactions, housing relocation records and medical records in various fields to provide an adaptive understanding of other issues.

4.5.3 Key insights for Greater London

A significant practical implication of this research is the empirical findings of Londoners' travel behaviours in recent years and under the impacts of the emergent global health crisis in early 2020. Linking clustering results with their socio-demographic variables allows policymakers and service providers to learn activity-travel patterns of different cohorts, abrupt changes when facing a crisis, and potential transition directions in the existing regime. The transport authority (i.e., TfL) can adjust the transport planning and services adaptively in London based on robust evidence.

This chapter first investigated the urban mobility in Greater London in a middle-long term, from 2015 to 2019. Then we assessed the early impact of the COVID-19 pandemic, which is a short-term analysis. The empirical study provides both short-term and long-term insights into the activity-travel patterns in Greater London.

Our empirical findings show geographical and socio-demographic differences in Londoners' activity-travel behaviours. For the two youths groups on weekdays, more youths attend schools/colleges in Outer London. Youths in Inner London are more ethnically diverse, older and from lower income households. For the workers who work in mixed places, more workers live in Outer London and work for longer hours. They tend to be older and whiter, compared to mixed places day workers living in Inner London. For people who both work and live in Inner London or Outer London, the Outer London workers are more likely to be less wealthy, middle-aged, males and from BAME groups. For non-workers, Inner London stay-at-homes tend to have lower incomes and are from more diverse backgrounds.

The emerging results in early 2020 show that, although major changes in travel behaviours happened after the legal restrictions, there were already some self-organising changes occurring before governmental interventions. Facing the health emergency, the self-organising behaviours in London included reducing public transport use, increasing time at home, and more shopping activities at weekends. The early dynamics in behavioural changes has been captured. The findings can help TfL understand changing behaviours of travellers and adjust transport services adaptively.

4.5.4 Implications for smart transport governance

As mentioned in chapter 2, travel pattern is an important area in smart transport management. Daily activity patterns and emerging changes are crucial for transport in smart cities (Kandt and Batty, 2021). The empirical findings can further support smart transport development and governance in Greater London. Greater London ranks first in smart transport, with the best performances in public and private transport. The evaluation result in chapter 3 shows that London ranked first in its accessibility and innovation aspects while environmental sustainability is a weakness for London. Smart products such as bike-sharing and smart policy/management that aim to build “healthy streets” as well as smart data analytics have the potential to help enhance sustainability in London (Moscholidou and Pangbourne, 2019).

This chapter provides a method for smart data analytics in London that can support transport interventions such as night-time economy, smart commuting, and healthy street. For instance, London’s main transport strategy, Mayor’s Transport Strategy 2018, aims to improve accessibility, sustainability, and innovation, with approaches of healthy streets. To improve healthy streets, active, inclusive, and safe travels are advocated, aiming for 80% of trips in London to be made in sustainable modes (walking, cycling and public transport) by 2041. We identified the main modes of each group and the socio-demographic profiles of these travellers. At weekends, the recreation groups mainly used private transport as the travelling means. The pandemic has further amplified the use of private transport. Thus, the population in these groups can be targeted in the post-COVID19 smart transport transitions. The empirical study provides an overall insight into different travellers in London. To support specific smart transport projects such as smart commuting, further study can zoom into target groups (e.g., workers' groups).

4.5.5 Limitations and further directions

This analysis is not without limitations. First, the dataset we used in this study contains socio-demography and trip information from 2015 to March 2020. Although we analysed the early impact of COVID-19 by mining the sequential data from January to March 2020, more data from the LTDS survey or other sources is needed to supply a more robust understanding of the impacts of the pandemic on travel patterns in London. A further direction is to analyse the LTDS data after March 2020 when it is available or to use data from other sources as a supplement to generate a comprehensive knowledge of the pandemic influences on the transport sector in London. Secondly, the LTDS dataset in this study is often seen as traditional "small data". In the big data era, results from "small data" are often critiqued for lacking detailed information on finer spatial or temporal scales. Thus, another further direction is to apply the same method to big datasets such as mobile phone big data to provide a more robust understanding of travel patterns. Lastly, we identified the main socio-demographic determinants in each cluster, but we did not develop predictive models. Another important direction in activity-travel pattern studies is to develop predictive models through the models we used in this study and machine learning/deep learning models with more predictors (e.g., built environment variables) in the future.

4.6 Conclusion

This chapter provided a new methodological approach to recognise daily activity-travel sequences and learn patterns of travellers' behaviours. An empirical study in London was presented, using LTDS data. Representative travel patterns are revealed through the fuzzy multi-channel sequence analysis. Eleven typical patterns have been identified on weekdays while eight representative sequences have been found on weekends. The main determinants on weekdays are age, occupation, household income and car access while key factors on weekends are household structure, household income, accessible vehicles, and ethnic group.

Furthermore, the early impact of the COVID-19 pandemic has been revealed. The complexity of daily activity-travel patterns was reduced, and the main modes of trips changed. Although significant changes in travel patterns only occurred later due to legal restrictions, we could already see some small changes in citizens' behaviours, including the avoidance of public transport and the reduction of out-of-home activities. The mixed place workers' groups became larger, and an emerging pattern on weekends that contained new shopping activities was identified. Additionally, we found that working from home status, instead of a new determinant, has long been a key factor in travel behaviours on weekdays. The pandemic has amplified its impact.

Finally, the implications of adaptive thinking in understanding activity-travel patterns, the transferability of the proposed method, the key insights and policy implication of empirical results, as well as limitations and further directions were discussed. The findings can increase the understanding of complex travel-activity patterns, assist emergent pandemic control, and support smart transport governance.

Chapter 5 : Sensing impacts of COVID-19 on travel behaviours in London using social media data

5.1 Introduction

The COVID-19 pandemic has resulted in cities implementing lockdown measures, causing unprecedented disruption (e.g., school/shop/office closures) to urban life. Due to unprecedented restrictions on social gatherings and movement, cities have undoubtedly become the front lines in fighting the COVID-19 pandemic. The unexpected outbreak required actions from governments, citizens, companies, and other groups, which challenged the capacity of governance of cities and their sub-systems. Under this circumstance, the gap 4 of lacking ability to manage and plan for uncertainties (brought by COVID-19) should be addressed. From the lens of Complexity theory in cities, the initial emergence of the pandemic is unpredictable, and the impacts of different interventions cannot be known until the end of the pandemic (Haken et al., 2021). As mentioned in Chapter 2, what we can do amid the uncertainty is to develop robust analyses to better understand the emergence itself, potential impacts, and dynamic changes (Haken et al., 2021). Despite many cities have made efforts to adapt to the health emergence, adaptations of urban systems have been difficult (Haken et al., 2021). Managing the pandemic is, therefore, a complex urgent challenge to urban living and it needs smart governance with a comprehensive understanding of how the pandemic changes cities, urban subsystems, and behaviours of urban actors.

Transport is undoubtedly related to disease transmission within and across cities, and this sector is adversely influenced and needs governance (Zhang et al., 2020; Moslem et al., 2020; Batty, 2020). An urgent challenge is that emergency transport needs to continue to operate effectively during the pandemic. Another challenge in transport is that public transport is seen as high risk and people prefer to use private modes (Buhat et al., 2020; Pase et al., 2020). Different sub-systems have changed non-linearly and faced unique challenges and opportunities. To support a holistic understanding of fast and slow dynamics in transport systems during the pandemic, this chapter focuses on changes and opportunities in the transport systems. Specifically, we investigate the four main subsystems, namely private transport, active transport, public transport, and emergency transport in this chapter, using theory-inform and data-driven knowledge discovery, to support adaptive governance.

When facing such an unexpected abruption, the public has shown some self-organising behaviours in travelling even before strong restrictions (see chapter 4) and other behavioural changes due to mobility-related restrictions. In the process of mitigating the abruption, emerging public behaviours are crucial for altering transport service and planning, in which public opinion plays a vital role. On the one hand, the public response directly reflects immediate perceptions and expectations on transport-related issues (Blumer, 1948; Wlezien, 2017). On the other hand, it can be a valuable source of information that policymakers can utilise to adapt current measures and policies. In the era of big data, social media has emerged as a major source to sense public opinion, providing unique opportunities for supporting urban management from the bottom-up (Feezell, 2018; Hochtl et al., 2016).

Greater London has been strongly hit, with the highest mortality rate in both the first and second waves of the pandemic (Adam, 2021; Office for National Statistics, 2021). London has selectively deployed smart approaches to support emergency management. UDS has played an important role in crisis management in London (Smart London, 2020). To manage unexpected changes in the transport sector, a mobility report was created on 2nd April 2020, using data from Apple, Google and TfL (GLA, 2020). However, timely and transparent social media big data hasn't been deployed in smart transport governance. Thus, there was a missing opportunity in extracting quick intelligence through social media big data to support smart governance of crisis and transport in London.

This final part of the PhD research aims to discover public opinion towards the COVID-19 pandemic and urban transport within a global health crisis context; resultantly, this chapter provides critical and timely insights by which to understand and facilitate the crisis and transport management as well as post-pandemic recovery in a smart city with the help of big data. This chapter also fills in the methodological gaps of lacking the use of citizen-centric data and correct understanding of citizens' needs through mining crowdsourced social media big data and advanced text mining techniques.

The chapter is organised as follows. Section 5.2 analyses a concise literature review on the COVID-19 pandemic impacts on travel behaviours and management, the role of social media big data, and transport management in Greater London. Section 5.3 explains the methodology employed herein, including data collection, data processing, sentiment analysis and topic modelling. Section 5.4 presents the empirical findings. Section 5.5 further discusses the results and implications. Section 5.6 summarises the key findings and concludes the chapter.

5.2 Literature review

5.2.1 The impacts of COVID-19 on travel behaviours and governance

COVID-19 has largely changed the travel behaviours and activity patterns of many individuals (Buehler and Pucher, 2021). Many alternative activities and travelling approaches have emerged and rapidly diffused, including localised trips, virtual mobility, and increasing usage of private and active modes of travelling (Loorbach et al., 2021). From the multi-level perspective of transition, urban mobility as an important functional regime has been rapidly altered by mobility-restricted measures globally and locally (Loorbach et al., 2021).

The pandemic has triggered non-linear changes in different sub-systems and population groups in cities. In most cities, traffic volumes decreased in most transport sub-systems, and many people reduced their out-of-home activities and trips (Kolarova et al., 2021; Bari et al., 2021; Zhang et al., 2021). Public transport has been avoided and active transport has become popular (Buehler and Pucher, 2021; Das et al., 2021). Many public transport users increased their usage of cars while some car users shifted their main mode to active transport after the outbreak of COVID-19 (De Haas et al., 2020; Dingil and Esztergar-Kiss, 2021; Budd and Ison, 2020).

Due to mobility-related restrictions and increasing virtual activities, many individuals have been provided with chances to try alternative travelling modes and activities. The pandemic is thus a “window of opportunity” to change the unsustainable habitual behaviours to more sustainable mobilities such as virtual, electric, localised, and low-carbon mobility (Budd and Ison, 2020; Kanda and Kivimaa, 2020). The extent of behavioural changes is based on both global changes of pandemic and local conditions such as local governmental measures and socio-economic characteristics. For example, cycling activities generally increased worldwide, but the time and volume were highly influenced by local restrictions (Buehler and Pucher, 2021). Thus, it is necessary to investigate dynamic changes in different transport sub-systems locally.

Facing the global health emergency, transport governance needs to adapt to complex and dynamic changes, considering both risks and opportunities. The short-term changes can have long-term effects on social norms and practices (i.e., behavioural routines), which can stimulate socio-technical transitions in the existing mobility system (Sovacool et al., 2020; Loorbach et al., 2021). From a socio-technical transition perspective, COVID-19 is a meta-transition event at the landscape level that permeates all socio-technical regimes (Wells et al., 2020). Combining pressures from the macro landscape (i.e., pandemics) and niche-driven improvements (local transport governance) can allow some niche innovations and actors to penetrate the dynamic stable regime (i.e., existing mobility systems) and even disrupt or displace them (Argyriou and Barry, 2021). During the transitions, undesired practices, technologies and cultures (e.g., unsustainable modes of travel) can be phased out (Loorbach et al., 2021). Thus, it is necessary to identify the opportunities during uncertain transitions to support smart transport governance.

However, transport governance has become increasingly complex and policymakers are unsure about future mobility demands (Marsden and Docherty, 2021). During COVID-19, questions such as what could be better done and how can we better respond have been raised. The future of mobility has been debated (Budd and Ison, 2020; Laverty et al., 2020). Researchers made various suggestions, including adaptive planning that accepts uncertainties, collaborative governance that involves the public, adoption of technological tools to support fast responses, and integration of resilience concepts in governance (Haken et al., 2021; Shaw et al., 2020; Yang, 2020; Mao, 2020; Joyce, 2021). As mentioned in Chapter 2, non-linear changes in the urban and transport systems are unpredictable. We can only conduct robust analysis that is “good enough” to understand complex situations, drawing insights from different data sources, advanced data mining tools, and domain knowledge.

5.2.2 Social media big data in managing COVID-19 and supporting transport planning

Within UDS, social media has become a vital data source to support planning and management, providing comprehensive and real-time information generated by people. Social media data record human activities in cities with time, location, tags, texts, images and profile information, allowing researchers to explore many aspects of urban management (Niu and Silva, 2020). Facing emergent crises, social media can sense the fast dynamics and support quick responses. During COVID-19, researchers have used social media data to analyse information/knowledge dissemination (Chan et al., 2020), effective communication and misinformation (Cinelli et al., 2020; Gottlieb and Dyer, 2020), pandemic trends (Lu and Zhang, 2020), general public concerns (Abd-Alrazaq et al., 2020), and policy measures (Ahmed et al., 2020a; Samuel et al., 2020; Chen et al., 2022).

In transport governance, social media has proven its usefulness in providing new perspectives for the operation and management of transport systems via real-time big data and analytics (Politis et al., 2021; Nikitas et al., 2020). It has great potential in exploring public transport mode demands, assessing service qualities, sensing users' opinions, detecting accidents and abnormal events, enhancing communications between authorities and the public, and improving public engagement (Grant-Muller et al., 2015; Nikolaidou and Papaioannou, 2018; Cottrill et al., 2017). Regarding the COVID-19 pandemic, social media data can support related transport studies in sensing public opinion, identifying emerging behavioural changes, and increasing public participation.

Although social media platforms can provide real-time and massive user-generated content, mining public responses to transport governance issues still faces many challenges. First, social media data contains vast amounts of irrelevant information and how to select datasets for analyses of specific policy measures requires a well-designed process of data pre-processing. Second, posts on social media are unstructured textual data. Converting unstructured text of social media posts into insights requires the implementation of text mining techniques such as sentiment analysis and topic modelling. Sentiment analysis can reveal the positive, negative or neutral tones of content (Nielsen, 2011). Topic modelling is another widely utilised approach to extract main themes from unstructured documents (Isoaho et al., 2021). Although these text mining methods have been separately used in previous studies, how to integrate those methods in sensing public responses to dynamically changing transport systems is still lacking. Thus, we propose a robust methodological framework for monitoring public responses from social media raw data.

5.2.3 COVID-19 and transport governance in Greater London

The coronavirus first reached the UK on 31st January 2020, and the date 12th February 2020 witnessed the first case in London (Ghosh et al., 2020). To mitigate COVID-19, the UK central government and the GLA have put forward a set of mobility-related restrictions since 16th March 2020 (Hadjidemetriou et al., 2020). The GLA suggested London workers work from home on 17th March 2020. Schools and pubs were closed on 21st March 2020, and the city entered the first lockdown on 24th March 2020 (GLA, 2020). More restrictions on safer transport services were implemented but the public acceptance of these measures was unknown (Budd and Ison, 2020).

During the first lockdown, transport usage reduced to 10% of pre-pandemic status and home-working increased to 130% (GLA, 2020). At the beginning of the first lockdown, private transport decreased to less than 50% of the usual level and public transport was reduced by over 80% (Drummond, 2021). Cycling activities remained stable on weekdays and even increased on weekends. Walking activities increased during the lockdown (TfL, 2020e; Drummond, 2021).

The first lockdown was gradually eased after May 2020. Workers can partly return to offices from 10th May 2020 and schools started to reopen on 1st June 2020. The city then generally reopened on 4th July 2020. However, the city then entered the second and third lockdown on 5th November 2020 and 5th January 2021 respectively (GLA, 2020). As the first lockdown has disrupted the transport system and changed people's daily behaviours most, we focused on the time period of the first lockdown and reopen stage in this chapter.

During the pandemic, TfL has adapted its public transport service provision and operation (TfL, 2021). Transport services have been reduced since 16th March 2020 (Hadjidemetriou et al., 2020). TfL has grasped the opportunity to advocate safe and active travel during a time of uncertainty (Budd and Ison, 2020). Interventions such as the Low Traffic Neighbourhoods and the London Streetspace were introduced to support active transport (Aldred and Goodman, 2020). The Streetspace for London programme, including longer trial cycle lanes and low traffic neighbourhoods, started during the pandemic (TfL, 2021). Additionally, the Congestion Charge was extended to longer hours in seven days a week to avoid car-led recovery after reopening the city (TfL, 2021).

Transport governance has accepted the uncertainties in scenario-based planning. TfL adopted adaptive planning with five plausible futures with different possibilities, including the most optimistic to pessimistic scenarios (TfL, 2021). Nevertheless, a new analytical framework for tracking development in the transport sector is needed (TfL, 2021).

UDS plays a key role in urban and transport governance in London. London has utilised multi-sourced data during the pandemic. London has released data and presented analytical results on the London Datastore, an open access data platform (GLA and London Office of Technology & Innovation, 2020). For example, the COVID-19 Mobility report is the first webpage created on 2nd April 2020 to illustrate immediate impacts on the transport sector, using data from Apple, Google and TfL (GLA, 2020). Nevertheless, the GLA stated that it did not have good enough data and some existing data lacked detailed information (GLA and London Office of Technology & Innovation, 2020). The timely and transparent social media big data has not been deployed to support the emergent governance in GLA. It has the potential to provide insights into behavioural changes.

5.3 Research design

5.3.1 Research framework

During the first lockdown, Twitter users posted a large number of tweets within Greater London, discussing their concerns about the disease, governmental responses, measures etc. This results in a vast and valuable dataset that contains valuable insights from the wider public. In this chapter, we use Twitter data to sense public opinion and link the findings with official surveys of the LTDS (from Chapter 4) and the London COVID-19 online diary.

To understand public perception towards transport-related issues, we investigated opinion in four main subsystems, namely private transport, active transport, public transport, and emergency transport. A two-step approach was developed (Figure 5-1).

1) Detecting trends of sentiment polarity during the first lockdown and reopening.

This step first measured the daily trends of sentiment polarity to detect the general sentiment in London per day during the first lockdown and reopening. To detect the sentiment changes related to transport issues, two sets of Twitter data were selected.

2) Extracting topics from discussions about different transport subsystems

To further understand the public responses to different transport systems, we subset the dataset of COVID-19 related tweets discussing transport-related topics. Four subsets of tweets were filtered by relevant keywords. By applying dynamic topic modelling techniques, we extracted specific topics from each subset of reopening tweets and visualised the results.

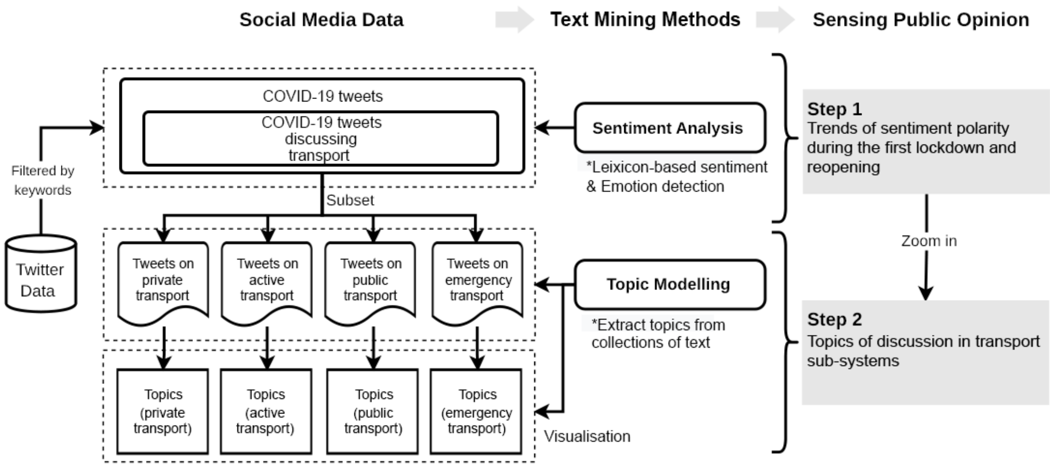

Figure 5-1: Analytical framework for Chapter 5

5.3.2 Twitter data collection and pre-processing

The social media data analysed in this chapter were collected through Twitter's developer Application Programming Interface, using the R package *rtweet* (Kearney, 2020). We collected relevant tweets posted in Greater London from April to August 2020. We conducted a geo search by setting the given radius at 20 miles, with the point of origin set as the central coordinates of London, enabling us to cover the administrative boundary of Greater London. The geo searches first found all tweets which were located within the geocode and then filter these tweets by language and keywords. English was set to limit the Twitter feed language. The keywords "COVID19 OR coronavirus* OR COVID" were set to collect COVID-19 related posts. The collected tweets include geotagged and non-geotagged posts.

As for transport-related posts, we further applied related keywords to filter tweets from the collected data. The detailed keywords and numbers of tweets in each group are listed in Table 5-1.

For each group, we first manually removed tweets published by bot accounts and news accounts (which could be identified based on the high frequency of posts and self-declared account descriptions). We calculated the dominant users who tweeted posts frequently and further analysed the top 50 users' accounts by frequency. By reviewing the users' descriptions and other profile information (e.g., username and profile image), we removed bot accounts and news accounts from the dataset.

Having finished the user screening,, we further cleaned the data through the Python package *NLTK* (Loper and Bird, 2002). We cleaned unstructured tweets by removing URLs, mentions, noisy words (i.e., Re-Tweets), newlines and extra whitespaces. Then, stop words such as 'the', 'that', and 'on' were removed. Finally, each word is converted into its base form.

Table 5-1: Keywords and numbers of collected tweets

	Groups	Subgroups	Keywords	Numbers of tweets
COVID-19	**All**		covid19, coronavirus*, covid	12,531,540
	Transport	**Transport**	transport, transportation, traffic, car*, journey*, trip*, travel*, mobility*, vehicle*, ambulance*, congestion*, parking*, station*, road*, railway*, motorway*, shuttle*, bus*, tube*, underground*, metro*, train*, rail*, tram*, ferry*, oyster card*, bicycle*, bike*, ebike*, scooter*, escooter*, cycle*, cycling*, walk*, walking*, on foot*, taxi*, cab*, uber*, passenger*	547,213
		Private transport	car*, bicycle*, bike*, ebike*, scooter*, escooter*, cycle*, cycling*, walk*, walking*, on foot*	149,507
		Active transport	bicycle*, bike*, ebike*, scooter*, escooter*, cycle*, cycling*, walk*, walking*, on foot*	89,303
		Public transport	bus*, tube*, underground*, metro*, train*, rail*, tram*, ferry*, oyster card*, taxi*, cab*, uber*, passenger*	166,291
		Emergency transport	ambulance*	22,434

5.3.3 Text mining methods

5.3.3.1 Sentiment Analysis

We used two sentiment analyses and a dynamic topic modelling technique to extract key insights from textual data in this study. Sentiment analysis is a popular approach to exploring insights from social media data. After cleaning all the collected tweets, we conducted sentiment analysis via the Python package *AFINN 0.1* (Nielsen, 2011). This is one of the fastest and most used sentiment analysis tools and has been broadly applied. For instance, previous urban studies applied this method in exploring public sentiment toward urban phenomena and urban planning measures (Chen et al., 2020; Hollander and Renski, 2017). The AFINN sentiment analysis is a lexicon-based method to score each word by comparing it to the scores of an existing English word list (Al-Shabi, 2020). After scoring each word in the tweets, the result of a post sums up all scores in a sentence. Each word has a score between -5 to 5. A negative score denotes a negative sentiment while a positive result denotes a positive sentiment, with zero meaning neutral sentiment (Nielsen, 2011). After generating the scores for all tweets, public attitudes and emotions can be revealed. For all groups (see Table 5-1), we calculated the means of everyday tweets to provide daily sentiment trends. The sentiment results in the COVID-19 group and transport group thus revealing public feelings, emotions, and sentiment towards the pandemic and different transport subsystems.

The second method is emotion detection analysis, using the Profile of Mood States (POMS). POMS is widely used to measure typical moods in clinical and social psychology (Petrowski et al., 2021). POMS is a psychological rating method to assess distinct mood stages of anger, depression, fatigue, vigour, tension, and confusion. Anger refers to the mood of anguish and hostility while depression contains the feeling of sadness, loneliness, guilt, worthlessness, and hopelessness. Fatigue is associated with low energy and inertia while vigour is linked with the cheerful mood. Tension is the status caused by anxiety and impatience. Confusion is characterised by bewilderment and cognitive inefficiency (Lin et al., 2014). We applied a pre-trained recurrent neural network model by Colneric and Demsar (2020) to classify the six mood states of posts.

5.3.3.2 Topic Modelling

Topic modelling is one of the most powerful text mining tools for exploring semantic structure from a collection of texts. Latent Dirichlet Allocation (LDA), as a generative probabilistic model, is commonly used to extract topics from a collection of documents (Capela and Ramirez-Marquez, 2019; Taecharungroj and Mathayomchan, 2020). In the LDA topic model, the collection of documents is referred to as the corpus; items within the corpus are referred to as the document, with specific words in documents called terms. The LDA model assumes that a document is generated according to the following process: 1) decide the number of words N that the document will include before randomly choosing a distribution over topics and 2) generate each word in the document. In step two, the model probabilistically draws one of the K topics according to the distribution over topics sampled above, and probabilistically draws one of the words according to the topic's multinomial distribution. Based on this generative model, the LDA model backtracks from the documents to discover the topics that are likely to have produced the corpus. The LDA topic model assumes that each document in the corpus is a mixture of K topics that are characterised by terms with certain probabilities. Each latent topic is characterised by a distribution over a fixed vocabulary. The details of LDA can be found in Blei et al. (2003).

In this analysis, LDA topic modelling is utilised to reveal latent topics inherent to the data, with probabilities of terms from the documents, which are tweets discussing the transport systems. In the process of training the LDA model, we first created a dictionary representation for all tweets. Additionally, we removed rare terms based on their term-document frequency. Then, we transformed the tweets into a vectorised form with the bag-of-words representation. These vectorised tweets were input for LDA training. The output was a list of topics with probabilities ascribed to each topic.

We used the R package *topicmodels* to conduct LDA (Grun and Hornik, 2011). We determined the optimal number of topics through the four metrics proposed by Griffiths and Steyvers (2004), Cao et al. (2009), Arun et al. (2010) and Deveaud et al. (2014), using the R package *ldatuning* (Nikita et al., 2020).

5.4 Results

5.4.1 Public attention and sentiment towards COVID-19

The pandemic has attracted great attention from Londoners on Twitter (see red line in Figure 5-2). “COVID-19” was one of the trending topics during the first national lockdown period. Under this circumstance, most of the citizens using social media have been alerted. The first lockdown was implemented at the end of March 2020. At the beginning of the first lockdown (in April), around 155,000 posts were discussing COVID-19 daily. May witnessed about 120,000 tweets every day. As time went by, people got used to and were bored with the pandemic-related information. Thus, fewer posts were seen in later months. The average number of tweets in June, July and August are 72,000, 55,000, and 41,000.

Regarding transport-related issues, they were not hot topics during the pandemic (see blue line in Figure 5-2). The daily discussion was between 527 to 17,234 each day from April to August. There were around 5,800 tweets in April each day. The daily discussion in May was about 6,000. Fewer discussions were seen in June, with about 2,300 per day. In July, particularly after the city was reopened, more debates on transport systems were found (around 3,400 per day). In August, fewer people discussed transport issues (about 1,500 every day).

Facing the pandemic, social media users were expressing negative sentiments all the time (see the red line in Figure 5-3). The average sentiment was between -4 to 0 each day. The most negative sentiment can be found on 14th June 2020, when people complained about the UK government for failing to mitigate the crisis as the UK had the worst death rate in European countries. Posts in July witnessed the most positive attitudes. Although the sentiment remained negative-neutral, flattening the curve, and reopening the city seemed to make people happier. However, more negative views were expressed in August. This can be explained by increasing daily cases. People become disappointed about the situation and governmental measures.

For transport-related tweets, the sentiment fluctuated, ranging from -5 to 3 (see blue line in Figure 5-3). Compared to the general COVID-19 discussion, we can find more positive discussions every month. The positive tweets are associated with thanks for frontline transport workers, support for active transport schemes and healthy street projects, the pleasure to go out (mainly in July and August) after a long time of mobility restrictions and the good news on vaccine development.

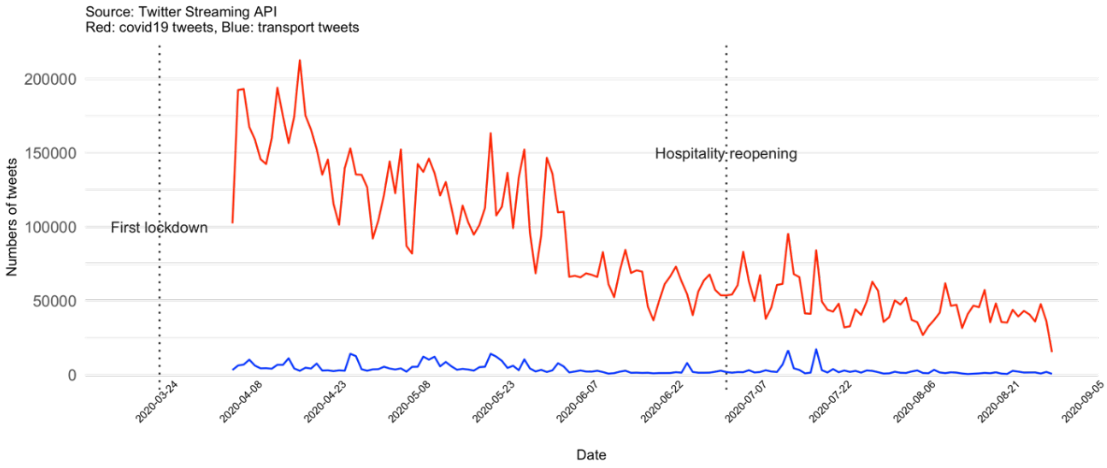

Figure 5-2: Numbers of tweets in Greater London (April to August 2020)

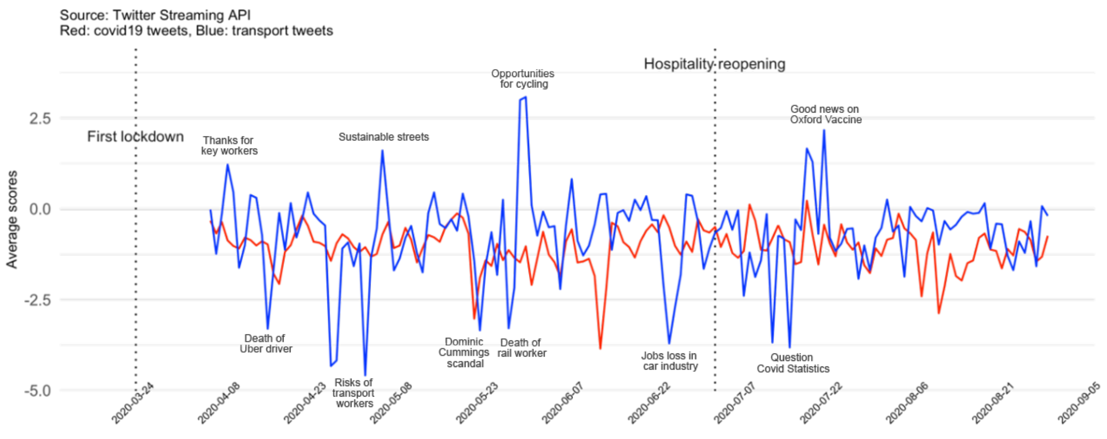

Figure 5-3: Overview of the sentiment of tweets in London (April to August 2020)

Regarding the emotional states of Twitter discussion, people extensively expressed their depression. In these discussions, people felt desperate when losing friends, terrified when hate crimes occurred, and disappointed at bad policies that failed to mitigate the pandemic. Apart from depression, anger and confusion were widely articulated (see Figure 5-4). In anger posts, people were annoyed by the lack of personal protective equipment, ventilators, and other materials at the early stage. The anger posts also blamed the government for not clearing out cases before reopening. Among confusion posts, Londoners stated that they are uncertain about the future during the pandemic and in the post-pandemic world. Vigour emotion was the least seen in COVID-related posts and this sentiment was usually linked with vaccines, National Health Service (NHS) workers, and reopening measures.

Similar to the sentiment results, the emotion states varied in the transport-related discussion (see Figure 5-5). Although people also extensively expressed their depression towards the transport system during the pandemic, more vigorous tweets can be found when people discussed transport issues. Citizens showed positive attitudes toward active transport and cleaner air during COVID-19. For example, people were satisfied with the School Streets scheme for less pollution and safer roads in August. Additionally, Londoners were glad to go for a walk during the lockdown and happy to go out after the lockdown.

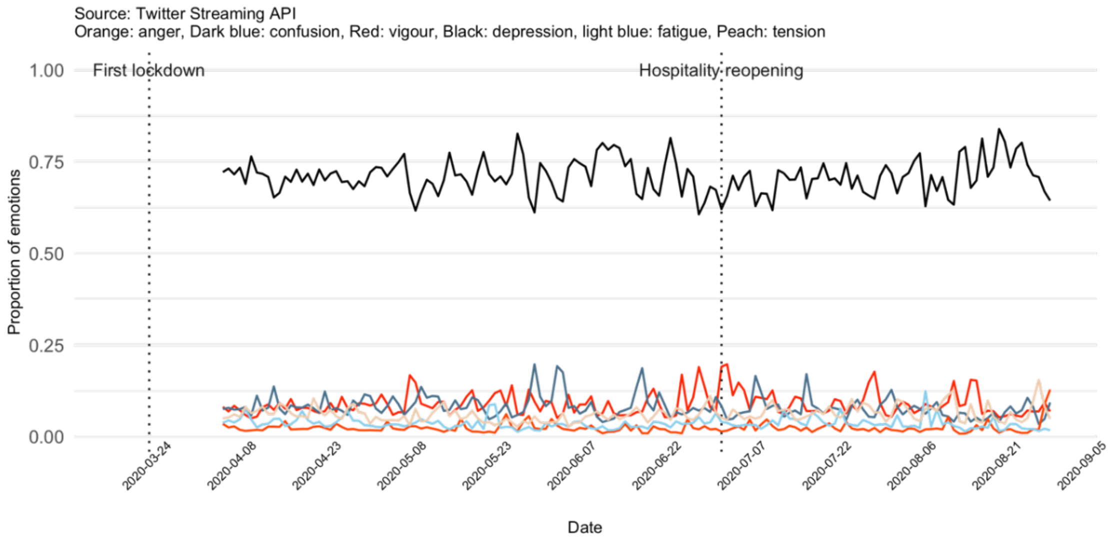

Figure 5-4: Emotion states of COVID-19 tweets

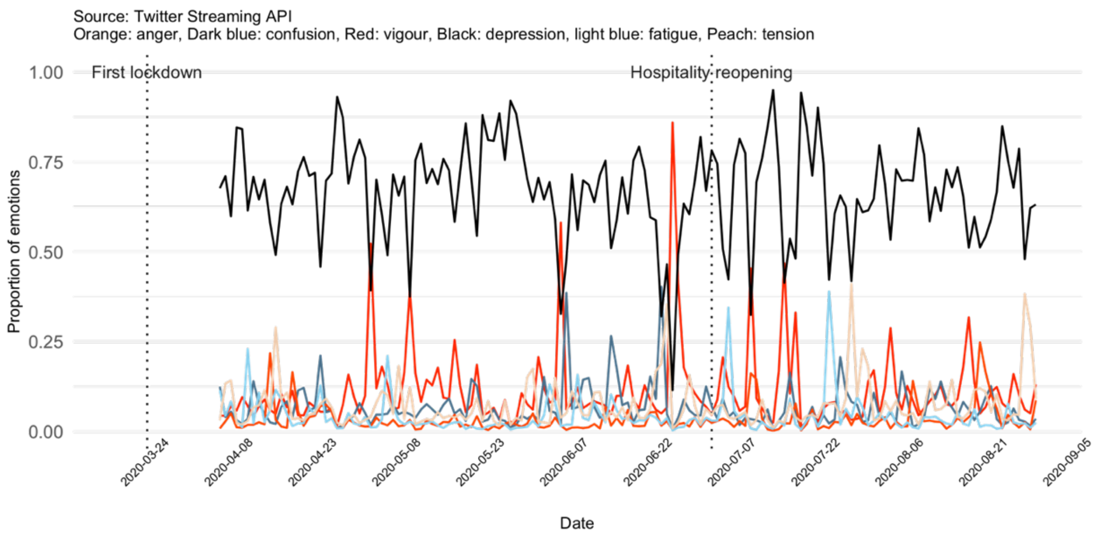

Figure 5-5: Emotion states of transport tweets

5.4.2 Main topics for transport sub-systems

As mentioned in 5.2.3, the pandemic has triggered non-linear changes in different transport sub-systems. The pandemic has changed the future directions of the transport sector and sub-systems in London (Campaign for Better Transport, 2020). To reveal detailed insights into changes in the transport sector during the pandemic, we explored dynamic latent topics in private transport, active transport, public transport, and emergency transport.

The topic modelling results are shown in Figure 5-6 and Table 5-2. In Figure 5-6, temporal changes of each topic within a transport subsystem are shown. For example, in the right bottom model for emergency transport, the proportion of the first topic (E#1) gradually increased from April to July and then decreased in August. The bar height indicates the importance of the topic in each month. Table 5-2 lists the most relevant terms for each topic. After reviewing key terms in each topic and linking them to relevant news and policies, we manually named all topics.

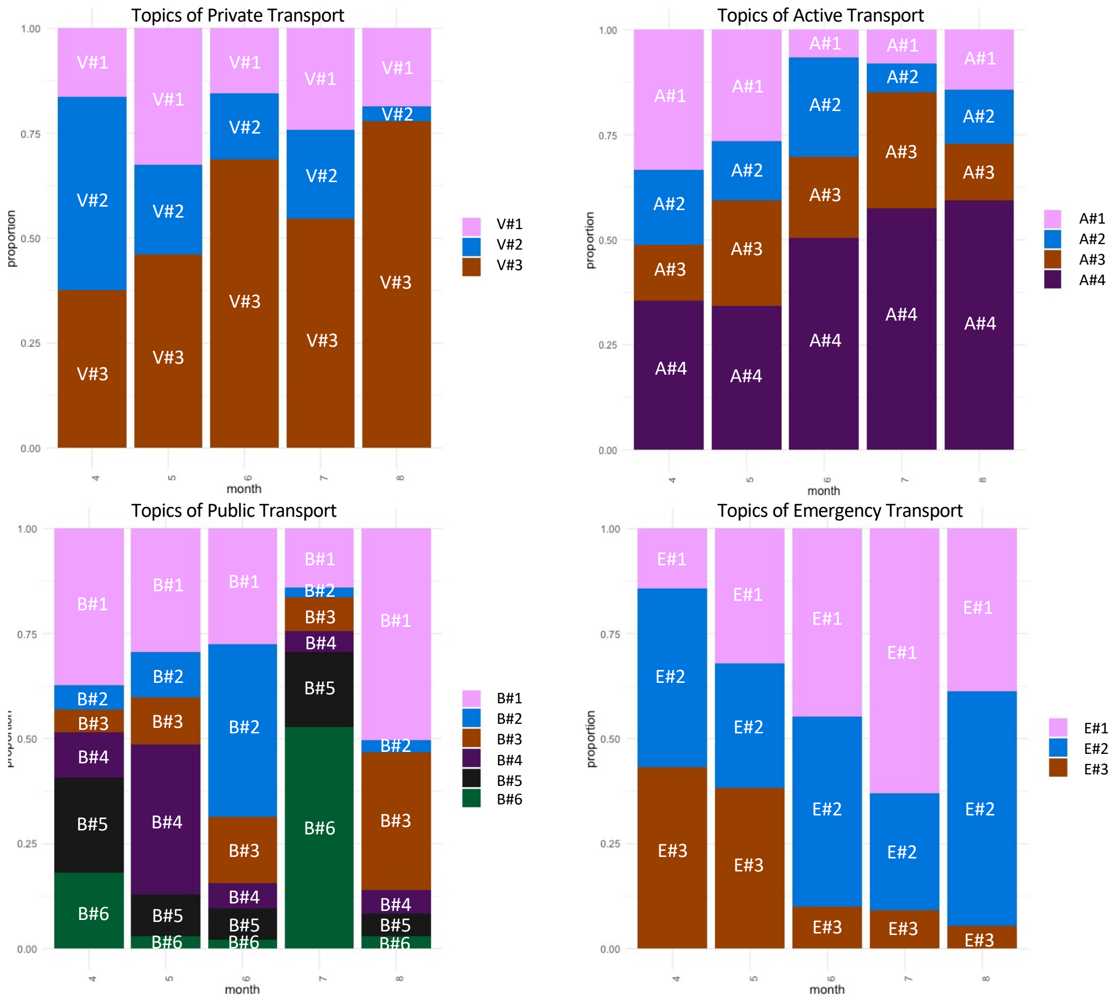

Figure 5-6: Topics of transport discussion

Table 5-2: Topics and key terms of transport discussion

System	Topic	Terms
Private transport	V#1 Governmental measures	Covid, TfL, car, coronavirus, London, risk, govern, city, walk, thing, time, NHS, free, worker, make
	V#2 News events	Landlord, car, die, require, day, pay, hours, lock, death, place, change, result, found, upfront, slept
	V#3 Active mode	Covid, walk, car, coronavirus, cycle, people, lockdown, bike, govt, public, time, live, London, transport, road, work, street, death, back, care
Active transport	A#1 Sponsored walk event	Walk, rais, garden, year old, Tom, Moore, NHS, lap, people, captain, veteran, war, coronavirus, street, covid
	A#2 Public criticisms	Walk, covid, govt, TfL, time, coronavirus, death, made, nation, year, hear, account, toll, critic, window
	A#3 Government plans	Covid, London, cycle, city, TfL, walk, travel, plan, transport, world, air, announce, increase, govern, improve
	A#4 Active commuting	Walk, covid, cycle, coronavirus, bike, people, lockdown, work, safe, delivery, day, transport, lane, distance, London
Public transport	B#1 Air travel	Covid, passenger, coronavirus, TfL, train, London, bus, govern, travel, airport, arrive, money, day, staff, quarantine
	B#2 Disappointing news	Covid, contract, PPE, Brexit, company, plane, bring, wrong, dear, paint, bankrupt, satellite, WW2, ventilators, Pest control
	B#3 Fare price	Covid, posit, train, test, bus, coronavirus, death, TfL, people, don't, remove, week, yesterday, put, fare
	B#4 Public transport workers	Work, coronavirus, lockdown, London, tube, back, bus, transport, pack, public, people, avoid, train, govt, passenger
	B#5 Safer travel restrictions	Die, driver, bus, mask, covid, transport, wear, public, people, live, told, lost, dad, save, left
	B#6 Vaccine and safe travel	Coronavirus, train, face, system, show, vaccine, tube, immune, key, development, trail, oxford, university, early, appear
Emergency transport	E#1 Financial assistance	TfL, covid, bailout, coronavirus, govern, Khan, staff, London, require, level, run, Sadiq, fare, mayor, finance
	E#2 Warnings of COVID-19	Ambulance, covid, coronavirus, TfL, die, hospital, worker, patient, support, service, Tahir, NHS, air, PPE, tragic
	E#3 Risks of workers	TfL, London, service, covid, ambulance, worker, thing, driver, cleaner, need, NHS, deal, emergence, staff, care

5.4.2.1 Private transport

Reducing the mode choice of private car usage is one of the main objectives in transport governance before the pandemic. During COVID-19, although road transport activities have generally been reduced, cars are preferred to minimise the risks of contact (Budd and Ison, 2020; Sung and Monschauer, 2020).

We found three topics in private transport discussions, as shown in Figure 5-6. Topic V#1 is associated with governmental measures to mitigate risks in the transport system. This topic also discusses the risks of travellers and suggestions for offering free travels for NHS workers. Topic V#2 is about special news events. The sad news of an Uber driver being evicted from home by the landlord and left to die in the car has raised much depression and anger (Nagesh, 2020). As active transport is included in private transport, topic V#3 especially links with active modes. People increasingly discussed active transport as time went by, showing more people were considering shifting modes.

5.4.2.2 Active transport

Active transport is one of the priorities in transport planning, aiming to increase sustainability (Campaign for Better Transport, 2020). As discussed in section 5.2.1, walking and cycling were chosen by many people during COVID-19.

From active transport discussions, we extracted four latent topics. Topic A#3 concerns new plans and announcements to support active travel. People discussed relevant plans for widening pavements, creating new cycle lanes, and converting road places into COVID-19 streets that support social distancing for pedestrians and cyclists. The announcements and measures that people cared about include measures and funds for reallocating road spaces, repairing bikes, and providing cycle parking spaces (DfT, 2020c), projects to transform roundabouts (TfL, 2020d), and the plan for safe and sustainable travel after reopening the city (TfL, 2020c).

Topic A#4 mainly discusses active commuting. “Cycling to work” was an increasingly discussed topic, particularly after the lockdown when people started to partially return to offices. This indicates that more people were considering walking or cycling to work, which can lead to long-term habitual changes in travel behaviour. This may be associated with alternative experiences during the pandemic and the government encouragement of “walk and cycle if you can” (TfL, 2020c).

Besides, topic A#2 calls for an emergency plan to support active mode and criticises the government for failing to mitigate the pandemic. Topic A#1 is related to the sponsored walk by Captain Tom Moore to raise money for NHS (BBC, 2020a).

5.4.2.3 Public transport

During the lockdown, a huge reduction in journeys was seen on London underground and buses, which led to a large amount of income loss of TfL (TfL, 2020a). The governance capability of TfL has been criticised on social media and news media (Mackintosh, 2020). Later, the financing support package (i.e., the “Funding Package”) has been provided to the TfL since 2020 (TfL, 2022).

Six topics are extracted from the public transport discussion. Topic B#2 is related to criticisms of the government. Disappointing news such as the lack of personal protective equipment for the TfL workers, the deaths of public transport workers, and overcrowding conditions in tubes have drawn much attention from the public (BBC, 2020b; Paton, 2020). Citizens were angry and sad about these situations in the transport system. The disappointing news can further lead to people’s avoidance of public transport. Topic B#4 particular concerns public transport workers such as bus drivers and taxi drivers. This topic has been widely discussed in the early stage, especially in May. Londoners felt sorry for losing many workers and urged the government to improve their working environments.

Topic B#5 is about COVID-19-related restrictions on public transport, especially the rules of wearing masks. Londoners were extensively concerned with restrictions in their posts in April. Topic B#3, a hot topic in the later stage (August), is associated with fare price. People were worried about raising fares due to the financial shock of public transport services. Citizens were also concerned about the financial situation of TfL.

Additionally, topic B#1 is related to air travel, including new restrictions on flights, and financial loss of passengers and airlines. The topic has been widely discussed in April, the early stage of the pandemic. The topic has also attracted increasing attention after the city was reopened. The last topic B#6 concerns vaccines and safe travel. People were discussing vaccines and wished to travel safely again. The topic was mainly seen in July.

5.4.2.4 Emergency transport

Three topics are identified in emergency transport. Topic E#3 concerns the risks of ambulance workers. News about deaths of ambulance staffs worried Londoners. This topic has been widely discussed in April and May. Topic E#1 is related to the financial assistance of ambulances and frontline workers. The public suggested providing financial supports for ambulance employees. This topic was widely discussed in June and July.

Lastly, topic E#2 is associated with warning messages that also mentioned “hospital” and “NHS”. The hashtag “ambulance” was widely used to raise people’s awareness to protect themselves. This topic can be seen in all months.

5.5 Discussion

5.5.1 Robust understanding through linking multi-sourced data analytics

In the result section, we identified quasi-real-time dynamics in transport systems during the pandemic. The real-time public perceptions can supply the key findings of official surveys and other sources of big data in the GLA COVID-19 Mobility Report, the GLA COVID-19 online diary and the LTDS.

The GLA COVID-19 Mobility Report has used multi-sourced data to measure different aspects of transport changes, providing direct quantitative information on activity and trip changes (GLA, 2020). However, the qualitative information on public acceptance, concerns, and underlying behavioural perceptions are not included. The findings from this chapter can provide qualitative insights into these aspects.

The opinion research team at GLA organised the COVID-19 online diary to capture the opinion of 20 citizens in London. The online survey started in mid-May 2020 and ended in mid-July 2020, which overlaps with the time period of this study (GLA Opinion Research, 2020e). However, the sample size was relatively small, and the results were first revealed in June 2020. Our findings can supply this survey to present opinions from a larger sample and in quasi-real time.

The LTDS provides a representative sample of London householders of around 8000 a year to show their travel-related information (TfL, 2020b). However, the official survey of LTDS has not been updated for external research use at the time of thesis writing (March 2022) and the data collection process (i.e., interview) has been severely impacted by COVID-19. Under this circumstance, mining social media big data can overcome the long update time of the conventional dataset and is not affected by the mobility-related restrictions in the data collection process. Our findings can provide quick knowledge discovery to supply the findings from the LTDS. The main results of the official surveys can be used to validate the findings from social media big data.

To validate our results, we compared our findings with the official surveys - London COVID-19 online diary and the LTDS-based findings. We first compared our results with the findings from London COVID-19 online diary. The report from weeks 1 and 2 (late May) states that Londoners experienced wide-ranging emotions. The main emotions were “fearful and anxious; coping and trying to make the most of it; frustrated and had enough; and financially insecure and stressed” (GLA Opinion Research, 2020a). Similarly, the emotions of tweets in May varied but mainly showed depression, confusion, and anger. It should be noted that the depression state in this analysis contains fearful feelings (Colneric and Demsar, 2020).

The weekly report from early-mid June shows that Londoners longed for economic recovery and cared about safety (GLA Opinion Research, 2020b). The safety-related concerns can be found in all transport sub-systems, including topic A#4 in active transport, B#5 in public transport and E#2 in emergency transport. Key findings from late June show that people started to get fatigued and “less engaged with the pandemic” in terms of seeking information (GLA Opinion Research, 2020c). This is accorded with our findings of decreasing public attention towards the pandemic over time.

The last weekly report from early-mid July illustrates that Londoners supported the idea of a 15-minute walking/cycling city but questioned relevant facilities such as cycle infrastructure (GLA Opinion Research, 2020d). This echoes the increasing attention towards active transport (see topic V#3 and topic A#4) and concerns about facilities to support active travel (see topic A#3).

We additionally compared our findings with the LTDS-based results in Chapter 4 and the TfL reports. An emerging trend of modal shift before the lockdown was identified in Chapter 4. In this chapter, we found that people increasingly discussed new modal choices after the lockdown. This may be because people were exposed to negative news about public transport and more discussion on walking and cycling. The information can result in a shift from public transport to active transport when the city reopened.

The findings in this chapter can be confirmed by the TfL report 14 on travel in London (TfL, 2021). Report 14 states that Londoners were worried about the safety issues of public transport. Londoners needed safe, reliable, and sustainable public transport with safety measures (TfL, 2021). The safety-related restrictions and measures were widely discussed in public transport sub-systems, as shown in topic B#5. Topic B#6 on vaccines and safe travel was corresponding to the citizens' needs for safer travel with effective measures. Additionally, Londoners felt sorry for the death of public service workers (TfL, 2021). This echoes our findings in topic B#4. Moreover, due to the worries about public transport, the report shows that Londoners were considered substitution types of travelling (TfL, 2021). The topics of V#3 and A#4 are related to the consideration of the alternative mode.

In summary, our findings can be validated by the official behavioural and travel survey. Compared to the official surveys, our findings from quasi-real-time social media big data have the potential to provide timely intelligence on dynamic changes. Combining the main findings from both big and small datasets can provide valuable knowledge to support a robust understanding of the potential mobility futures.

Timely and transparent social media data has a chance to assist in quick responses to the abrupt crisis and there were missing opportunities of grasping the crowd wisdom of citizens during the pandemic outbreak as the social media data was not included in the decision-making process. The temporal dynamics of public attention can be used as an opportunity for effective communications. The public attention was high in the first two months of the lockdown (April and May) and then faded. The early stage was an opportunity to effectively communicate with the public as more people were listening. With great interests, citizens can understand the problem in a short time period. Also, for mobility issues, there was another peak in discussion when the city was reopened in July, which could be a second opportunity for effective communication. Apart from mobility issues, the social media dataset can also provide quick insights for other important governance topics such as reopening (Chen et al., forthcoming) and vaccination (Puri et al., 2020).

5.5.2 Towards adaptive transport transitions

From the empirical findings, we identified positive attitudes towards the mobility system despite the general negative sentiment when people discussed the pandemic. Opportunities emerged amid the uncertainties brought by COVID-19. COVID-19 created a landscape shift and opened opportunities for niche innovations to emerge into the existing mainstream regime (Griffiths et al., 2021).

The pre-COVID smart transport novelties have been accelerated or hindered by the pandemic. For example, electric vehicles, especially private vehicles, were less influenced by the pandemic (Sovacool et al., 2020; Kanda and Kivimaa, 2020). The mobility-as-a-Service and shared mobility (carsharing) have been weakened as people were less willing to use public transport and shared vehicles (Sovacool et al., 2020; Kanda and Kivimaa, 2020). Bike sharing, on the contrary, has been first hindered and then accelerated after lockdown (Li et al., 2021). Among these smart transport niches, electric vehicles and bike-sharing are more likely to emerge into the existing regime during or after the pandemic.

When these niche innovations change the regime, they may not lead to sustainable mobility. Governance is needed to direct the transition pathways to socially desired directions (i.e., sustainable/inclusive/just mobility) (Sovacool et al., 2020). A citizen-centric adaptive approach can support complex transport governance, as illustrated in chapter 2. Citizens are the main users of public, private, and active transport in the post-COVID cities. Their acceptances of new niche novelties and more sustainable travel behaviours are crucial for transforming the existing mobility regime. Thus, citizens' needs and opinions should be highlighted in transport governance. Transport authorities and policymakers can make use of citizens' wisdom, so as to adaptively plan for socially desired futures (Griffiths et al., 2021).

Additionally, there should also be institutional changes within the local and transport authorities in terms of mindsets and skillsets. Amid uncertainties, policy makers and transport planners should see the black swan event from a perspective of complexity, making use of temporal and spatial dynamics, grasping opportunities from mining citizen-centric data, and adopting adaptive thinking in crisis management. Real-time social media big data, as citizen-centric data, is a key to governing with added value and supporting inclusive recovery. Real-time analytics can reveal the flexibility and alternative choices of citizens as well as the potential impacts on changing the existing regime.

As for London, TfL has already accepted uncertainties in future mobility and actively invented active transport during the pandemic. Nevertheless, social media data can further support smart governance in quick intelligence and rapid responses. The data results show the citizens' strong concerns on safety issues, worries of transport workers, and changing restrictions in the public transport system. In the long-term, to achieve a sustainable mobility future, the public transport system needs to be one of the main modes for Londoners. By identifying the underlying reasons hampering

Londoners' choice of public transport mode, TfL can further alter operational approaches and bring back public transport users. Active transport, as widely discussed by many Londoners, can be a new commuting mode for workers.

5.6 Conclusion

This chapter focuses on public opinion and sentiment in different transport sub-systems during the COVID-19 pandemic. Our findings show that public attention toward COVID-19 related discussions was high (right after the lockdown), but public attention decreased as time went by. Among the COVID-19 discussions, transport-related posts were in a small proportion (less than 5%). We found that Londoners increasingly discussed transport-related topics in the early stage after the national lockdown (April and May) and the transport topics regained new attention in July when the city was generally reopened. These periods of high public attention could be opportunities for effective government-citizen communication through social media platforms.

The sentiment results on transport topics fluctuated but were more positive than the general COVID-19 discussions. The positive tweets are associated with thanks for frontline transport workers, support for active transport schemes, and the pleasure to go out. The negative sentiment is linked with risks, sadly loss of frontline transport workers and unsatisfied sanitary conditions.

The topic modelling results further provide detailed insights into public opinion towards different transport sub-systems. People paid more attention to changes in the public transport system. Londoners were extensively concerned with restrictions and safety of transport workers in the early stage. In the later stage, people worried about fares and wished to travel safely again.

In the private transport system, citizens paid great attention to the active modes. Londoners showed increasing interest in commuting through walking and cycling. Regarding emergency transport, the risks of workers were widely discussed, and the hashtag “ambulance” was widely used to send warning messages to alert citizens.

The analysis is not without limitations. Social media data is often criticised for its data representativeness. British social media users tend to be younger, better educated and more liberal with higher attention to politics (Mellon and Prosser, 2017; Spielhofer et al., 2019). Nevertheless, previous studies show that around one third to half of users have communicated on social media during emergencies and the large amount of analysed posts can reveal public concerns to a large extent.

We see the potential of our findings and methodology. First, our study can contribute to the existing scientific understanding of COVID-19 and mobility transitions in post-COVID recovery. Advanced text mining methods were used to mine social media data and thereby facilitated understanding of public attitudes towards the pandemic and different transport sub-systems. Second, the key findings herein have the potential to assist the transport governance in the case area by enhancing public engagement and identifying emerging opportunities in the dynamics. Text-mining results revealed the main focuses, concerns, and preferences prevalent during the pandemic. Based on our findings, transport operations and policies can be adapted to better meet the public’s needs, creating conditions for more public desired mobility transitions.

Third, big data analytics can support citizen-centric adaptive governance. Social media data provides a large data sample and diverse perspectives in understanding the transport sector during the pandemic and in the aftermath. The dataset and advanced analytical tools allow researchers, policymakers, and governors to gain valuable, real-time insights. Combining findings from both big and "small" data, a robust understanding of transport systems in uncertainties can be better revealed. More adaptive, sustainable, and inclusive transport governance could then be built.

Chapter 6 : Conclusion and final remarks

6.1 Introduction

This chapter presents the overall discussion and conclusion of the thesis, highlighting the adaptive governance framework and key empirical findings. This chapter is composed of four main sections. We first revisit the main gaps and research questions. The second section discusses the theoretical framework of citizen-centric adaptive governance toward smart transport. The empirical findings in case studies are then discussed. In the penultimate section, limitations in this PhD research and possible future directions in which the current study could be extended further are presented. The last section concludes this study.

6.2 Revisiting research gaps and questions

This study has set out to understand complexities in smart transport governance from theoretical, methodological, and practical aspects. To achieve this overall objective, we assumed that CTC and UDS can support smart transport governance in theory and practice. Through a systematic literature review, this thesis found nine existing gaps in smart transport governance. The gaps can be categorised into theoretical, methodological, and governance aspects. The nine gaps are as follows.

- Gap 1: lacking theoretical understanding of complexity theory and data science in smart transport governance

- Gap 2: lacking sophisticated approaches to understanding complexity in dynamic transport systems

- Gap 3: lacking the use of citizen-centric data and correct understanding of citizens' needs

- Gap 4: lacking ability to manage and plan for uncertainties

To address the theoretical gaps, this thesis has conducted systematic literature reviews in Chapter 2 and has discussed specific elements within the conceptual framework in the discussion parts of Chapters 3-6. Chapter 2 addressed the first gaps by extracting the main implications of smart transport governance from CTC and UDS. Seven implications are extracted from CTC and five guidance are found from UDS. Chapter 2 further tackled the theoretical gap by proposing a holistic framework that integrates the existing smart governance framework and main implications from CTC and UDS. The new theoretical framework consists of multi-level perspectives, smart domains with dynamics in smart cities, data-driven knowledge discovery and outcomes with possibilities (Figure 2-3).

Building on the conceptual framework, the empirical chapters zoomed into specific issues in smart transport governance. In the discussion parts of the empirical chapters, the importance of adaptive governance has been emphasised. Chapter 3 suggested adaptive transport planning for managing emerging technological smart transport products. Chapter 4 highlighted the adaptive understanding of citizens' activity-travel patterns for making existing governance smarter. Chapter 5 presented a robust understanding through multi-sourced data analytics and pointed out citizen-centric adaptive transport transitions in the post-COVID city. The theoretical framework is further revisited and summarised in the next section. By doing so, Chapter 2 and the discussion parts in the Chapters 3-6 answered the first research question: "How to support smart city governance with new planning theory and scientific trends?" (Figure 1-1).

Building on the new theoretical framework established, Chapters 3, 4, and 5 conducted empirical studies, responding to all other gaps. Multi-sourced data and mixed methods were deployed to unfold the complexity of the smart transport system and reveal citizens' needs in case studies.

Chapter 3 first addressed the methodological gap 2 by reviewing existing indicators of smart transport and identifying important new indicators that have not yet been included in smart transport assessments, to provide a more sophisticated toolkit for smart transport interventions and investments (Table 3-3). A robust and up to date evaluation framework was built (Table 3-4, 3-5 and 3-6) and applied to ten Combined Authorities and GLA (i.e., eleven English metropolitan areas) (Figure 3-2). The empirical result shows that Greater London has the smartest transport system in English metropolitan areas. Chapter 3 additionally tackled the governance gap 4 by summarising the existing governance in eleven English metropolitan areas (Table 3-2), identifying the linkages between smart transport ranking and policy interventions, and suggesting adaptive transport planning for managing the uncertainties in future mobility systems. Thus, Chapter 3 answered the second research question "How to build an evaluation framework to show the topology of smart transport development?" (Figure 1-1).

Chapter 4 has unfolded the complexity of daily activity-travel patterns in London through data-driven knowledge discovery. Chapter 4 first addressed the methodological gaps 2 and 3 by presenting an innovative individual-based spatiotemporal pattern recognition method with the acceptance of uncertainties (Figure 4-1). Chapter 4 found representative spatiotemporal patterns of Londoners and their social-demographic characteristics. Chapter 4 also tackled the governance gaps 4 in Greater London by revealing fuzzy groups of travellers and extracting sequential changes after COVID-19. Chapter 4 answered the third research question "How to extract the daily travel activity patterns and explain the complexity of typical patterns?” (Figure 1-1).

Despite Chapter 4 extracting the early impacts of the pandemic, it should be noted that the results were found a year after the outbreak of COVID-19 (as the survey data of LTDS19/20 was released in 2021). Chapter 5, in response to the rapid need to understand the dynamic impacts of COVID-19 in real-time, has primarily addressed the governance gap 4, using the uncertainties brought by COVID-19. Meanwhile, Chapter 5 additionally tackled methodologies gaps by mining public opinions from citizen-centric social media big data. A two-step text mining approach was designed to measure daily trends of sentiment polarity and emotion states, as well as extract main topics about different transport subsystems in Chapter 5 (Figure 5-1). Chapter 5 sensed timely and transparent public attention, sentiment and discussion on COVID-19 and transport-related issues. Chapter 5 further discussed the importance of robust understanding supported by multi-sourced data analytics and citizen-centric adaptive transport governance towards different mobility transitions, to tackle the governance gaps. A small part of Chapter 4 and Chapter 5 answered the fourth research question “How to sense the dynamic changes in the transport sector during COVID-19?” (Figure 1-1).

6.3 Discussion of theoretical framework

This thesis has provided a theoretical framework for adaptive governance based on existing analytical frameworks, CTC, and UDS. In existing smart transport governance literature, the multi-level perspective is often used to understand the dynamic stabilities and changing patterns in the complex smart mobility transformation process (Geels, 2005: vi; Geels, 2020). The main elements in the multi-level perspective are: 1) macro-level landscape changes that put pressure on existing smart city and transport regime, 2) meso-level dominant smart transport regime in dynamic stability, 3) micro-level niches innovations that have the potential to shift regime into different directions, and 4) adaptive capacity as a key in successful non-linear transition management.

From complexity theory in cities, seven implications are extracted, which are: 1) accepting latent possibilities of uncertain futures, 2) exploiting opportunities in uncertainties, 3) understanding complex issues with interdisciplinary knowledge and approaches, 4) using holistic/robust analysis to support decision making, 5) increasing the responsiveness to changes through adaptive governance, 6) enhancing adaptivity in new institutional frameworks, and 7) adjusting governance to temporal and spatial dynamics.

Through the lens of UDS, this study finds five main guidance: 1) conducting robust analysis linking big data and “small data”, 2) putting special focus on citizenry science through mining human-generated data and better exploration of citizens’ needs, 3) understanding urban dynamics in different time and space scales, 4) generating interdisciplinary insights through combining data results with domain knowledge, 5) linking data with theories to better inform planning and governance.

The nexus of smart transport governance, complexity theory in cities, and UDS lie in theory-inform data-driven knowledge discovery and adaptive planning to understand the fast and slow dynamics with uncertainties in the transport sector. The data-informed results for adaptive understandings can support smart transport governance and even smart urban management with more elasticities and higher adaptive capacity. Integrated main notions and implications, an holistic framework for smart transport governance is proposed, including macro-level landscape, smart city regime, niche context, data-driven knowledge discovery, and outcomes with possibilities

(Figure 2-3).

This study has analysed the macro-level (global) pandemic and technological innovations, regime-level transport developments and travel demand before the pandemic, and the niche contexts of local governance and niche products amid COVID-19 (Figure 6-1).

In the existing smart transport regime, main sub-systems (public transport, private transport, active transport, and emergency transport) are interrelated in smart cities and governance can lead and manage all other domains. In a smart transport governance, adaptive capacity is highlighted in facing the dynamics both in short-term and long-term. This study particularly investigated the long-term dynamics of existing smart transport elements, activity-travel patterns, and potential middle-long-term impacts of COVID-19 on the existing transport regime.

The wider macro-level landscape of the smart city regime contains wider trends and crises such as urbanisation, climate change, technological innovations, and pandemics. This study analysed the impacts of technological advancement and sustainable goals on smart transport development in Chapter 3. As the outbreak of COVID-19 occurred halfway through the PhD study and it changed the existing mobility system, this thesis paid special attention to the influence of COVID-19 in Chapters 4 and 5.

The micro-level concerns the niche contexts that can change the existing regime and shift the current system towards new regimes. Previous studies have mainly analysed niche-level smart transport products such as CAVs, bike-sharing apps, and MaaS (Mladenovic and Haavisto, 2021; Manders et al., 2020). This study has additionally investigated the local socio-economic conditions, local institutional settings, and transport interventions as they all have a chance to transform the mobility systems. In Chapter 3, smart transport products, interventions, and institutional settings in eleven English metropolitan areas were analysed. The local condition of Greater London was further discussed in Chapters 4 and 5. Regarding COVID-19, the niche-level innovations of home-working, localised mobility, active modes, electric vehicles, MaaS, and car- and bike-sharing were discussed in Chapter 5. The pandemic created a landscape shift and opened opportunities for niche innovations to emerge into the existing mainstream regime. The pre-COVID smart transport novelties have been accelerated or hindered by the pandemic.

The next main part of this theoretical framework is data-driven knowledge discovery. To support smart transport governance, Chapter 2 highlighted the importance of linking the findings from big and "small" data, including domain theory in result interpretation, adopting fuzzy approaches to understand uncertainties, and conducting dynamic modelling. The empirical studies applied mixed methods to support knowledge discovery from urban big and "small" data. Chapter 3 mined multi-sourced data in the case study, using indicator analysis. The detailed indicator can show the development in each niche-level innovation and the synthetic indices are designed to represent the overall performance in key aspects. Using "small data", Chapter 4 provided an innovative method to identify long- and short-term activity-travel patterns of Londoners with a higher level of uncertainties to support adaptive understandings. Chapter 5 conducted big data analytics on social media posts, using novel text mining approaches, to extract important public opinion on the impacts of COVID-19 in quasi-real-time.

In the final part of this theoretical framework, the outcomes come with possibilities rather than certain results. In the adaptive governance framework, researchers and policymakers should accept the uncertainties in smart transport as the mobility system is constantly changing and the future is unpredictable. Thus, governance with adaptive mindsets should be adopted to understand and manage uncertainties. Despite that mobility futures are unpredictable, planners and policymakers can actively invent the future transitions by creating conditions to influence transport systems to increase the possibilities of sustainable and just directions, allowing co-evolutions towards socially desired futures of the wider public. During uncertainties, grasping the opportunities to influence the changing directions in the transport system is a key to adaptive governance, as illustrated in Chapters 3, 4, and 5.

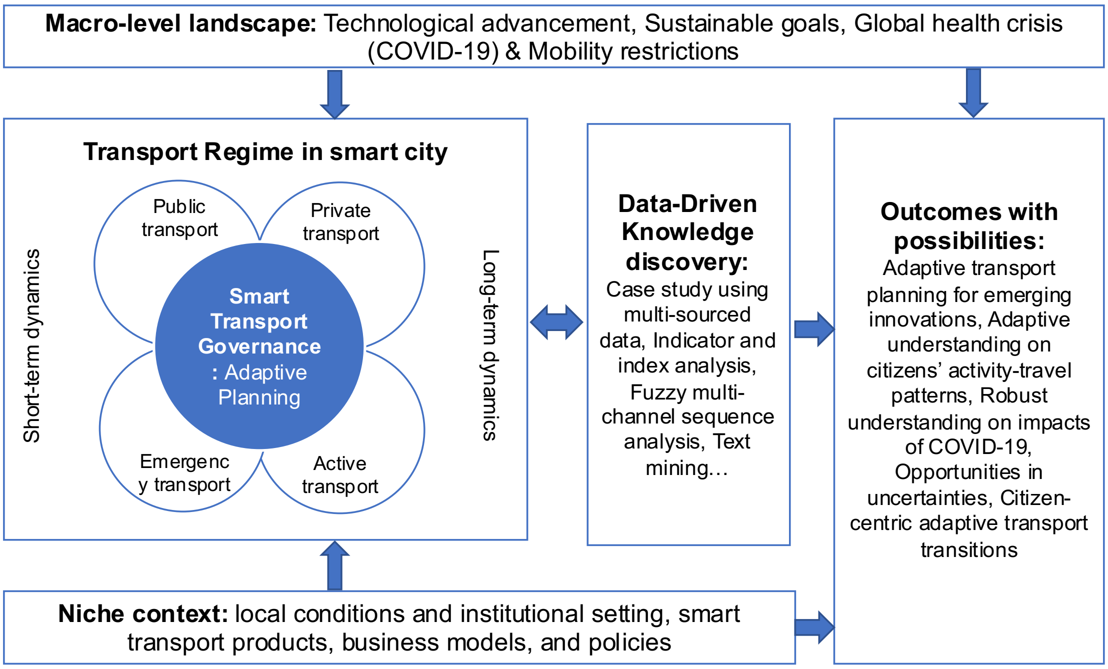

Figure 6-1: Empirical studies in the theoretical framework

6.4 Discussion of findings from empirical studies

This PhD study has conducted case studies in the English metropolitan areas, selecting ten Combined Authorities and Greater London as cases to provide a big picture of smart transport development in the UK.

In Chapter 3, a robust framework for assessing smart transport development was developed. The advantages and disadvantages of each CA and Greater London were presented in Chapter 3. The smartest transport can be found in Greater London, West Midlands, and West of England. Greater London ranked first in innovation and accessibility aspects while ranked worst in sustainability. In transport sub-systems, the strengths of Greater London are in the private transport (including walking and cycling) and public transport while its emergency transport system can be enhanced. Facing emerging innovations, adaptive planning is needed to prepare for uncertain future mobility transitions and transport authorities can actively invent future smart mobility towards inclusive and sustainable directions.

Chapters 4 and 5 conducted empirical studies in Greater London. Chapter 4 identified geographical and socio-demographic differences in Londoners' activity-travel patterns from 2015 to 2019. Eleven representative sequences on weekdays and eight typical patterns on the weekends, including stay-at-homes, day workers, and recreational groups, were identified. The eleven groups on weekdays are: 1) Outer London youths, 2) Inner London youths, 3) Mixed place day workers living in Outer London, 4) Mixed place day workers living in Inner London, 5) Outer London day workers, 6) Inner London day workers, 7) Inner London day workers living in Outer London, 8) Outer London stay-at-homes, 9) Inner London stay-at-homes, 10) Outer London afternoon recreation, and 11) External London mixed activities. The socio-demographic characteristics of Inner London and Outer London workers and non-workers groups varied. Key determinants are age, occupation, household income and car access.

The eight clusters on weekends are: 1) Outer London stay-at-homes, 2) Inner London stay-at-homes, 3) External London stay-at-homes, 4) Outer London afternoon-night recreation, 5) Outer London afternoon recreation, 6) Inner London afternoon-night recreation, 7) Outer London day workers, and 8) Inner London day workers. Key determinants are household structure, household income, accessible vehicles, and ethnic group.

Chapter 4 additionally identified the emerging self-organising changes when COVID-19 first hit London. The complexity of activities and travel reduced, with less time spent on education, work, recreation, and personal business and more trips by private transport. Representative sequences in early 2020 are slightly different from the pre-COVID clusters. More people were in the mixed places day workers groups on weekdays. A new cluster that contains people who went shopping in the afternoon after midday recreational activities emerged on weekends.

To supply quick insights on COVID-19, Chapter 5 further sensed the public opinion towards different transport sub-systems through real-time social media big data. Chapter 5 found that transport issues were not hot topics in COVID-19 related posts and the sentiment towards the transport sector fluctuated, with much depression and some vigorous emotions. The attention on transport first increased in the first two months after the national lockdown in late March and decreased in June when people get fatigued about the information on the pandemic. After the city was generally reopened in July, the transport-related posts increased again.

Chapter 5 revealed dynamic behavioural changes and opportunities in different transport sub-systems. More citizens considered walking or cycling to work. COVID-19 has provided an opportunity to promote active travel. Regarding public and emergency transport, risks and financial concerns were widely discussed. Governmental measures and recovery plans were also widely discussed in all transport sub-systems.

Chapter 5 identified missing opportunities for GLA and TfL to govern the transport systems smarter with rapid insights from the public. There was an opportunity for collaborative governance to support crisis management and transport governance during the pandemic.

6.5 Limitations and recommendations for future research

This PhD research has contributed to theoretical, methodological, and practical aspects of smart transport governance, as summarised above. However, the thesis has limitations worth noting and there are many ways to further improve the methods and results that the scope of this thesis did not allow.

First, more data is needed to refine the temporal and spatial scale in future studies. One limitation of this study was the time span of the data we used. The LTDS have been impacted by COVID-19 after March 2020. This thesis cannot access the LTDS data after March 2020 so the activity-travel changes of Londoners after the national lockdown in late March 2020 cannot be identified. Future studies can be conducted when the LTDS20/21 and LTDS21/22 become available. For social media data, this thesis only started to collect posts in April 2020 and decided to stop data collection in September 2020. Data before the outbreak of the pandemic (in late 2019 and early 2020) can allow further studies to identify public opinion in the early stage and use social media data to predict the outbreak. Additionally, Twitter posts in the second and third waves of COVID-19 can further allow us to find long-term changes. The temporal scale of the study can be further expanded.

Another limitation in the empirical studies was the spatial scale. Being limited to the spatial information in travel surveys, this study cannot reveal more detailed information on the spatial dynamics. Further study can make use of big data in finer spatial scales such as smart card data and mobile phone data. Additionally, this thesis focused on the metropolitan scale, the methods can be further applied to other spatial scales (from local communities to regions) to find the activity-travel patterns and potential travel behaviour changes spatially.

Secondly, more technical research is needed to improve the methodology for adaptive understanding. One limitation of the evaluation framework in Chapter 3 is that it lacks historical trends and cannot be automatically updated to show the temporal trends in the later years. A visual dashboard can be built to show the assessment framework with detailed indicators. Historical and new data can be further connected to the dashboard, allowing the dashboard to automatically refresh every year. Another limitation is that the sequence analysis is time-intensive in R. The codes can be rewritten in other programming languages such as Java or C to speed up the analysis.

Thirdly, this research has thrown up many questions in need of further investigation in the case areas. Considerably more work will need to be done in other English metropolitan areas to find opportunities in smart transport development. More investigation on detailed projects and policies in Greater London is needed to further understand the success story of smart transport and the role of adaptive planning in building the smart transport system in London. As the thesis paid more attention to the COVID-19 impacts in the later chapters, smart transport products such as Maas and electric vehicles have not been fully investigated. Further studies can provide more understanding of these smart products.

Lastly, the conceptual framework in this thesis has provided a method to support smart transport governance with new perspectives. Ideally, other researchers and practitioners will use this framework for planning and governing smart transport in the smart city and will also contribute to it. The theoretical framework can be applied to different cities in different countries because it is general enough that it is not anchored to the UK context. This study especially focused on smart urban transport in the smart city. The theoretical framework can also be applied to other interrelated sub-systems in a smart city such as smart environment and smart living.

6.6. Final Conclusion

In conclusion, this thesis has contributed to a deeper understanding of how adaptive planning and citizen-centric data contribute to smart transport governance in theory and practice. In theory, this thesis proposed a new theoretical framework, highlighting citizen-centric adaptive governance for smart transport. The empirical studies provided new evidence through data-driven knowledge discovery processes. Specifically, the study built a new evaluation framework of smart transport development in the English metropolitan areas, identified typical daily activity-travel sequences of Londoners and main socio-demographic determinants, and found the impacts of COVID-19 on different transport subsystems. The evidence can support adaptive transport planning towards sustainable and inclusive futures. Despite some limitations, this PhD research has contributed to existing theoretical and practical understandings of smart transport governance.

Reference

Aaditya B and Rahul TM (2021) A comprehensive analysis of the trip frequency behavior in COVID scenario. *Transportation Letters-the International Journal of Transportation Research* 13(5-6): 395-403.

Abbas K, Tawalbeh LA, Rafiq A, et al. (2021) Convergence of Blockchain and IoT for Secure Transportation Systems in Smart Cities. *Security and Communication Networks* 2021.

Abd-Alrazaq A, Alhuwail D, Househ M, et al. (2020) Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. *Journal of Medical Internet Research* 22(4).

Abu-Rayash A and Dincer I (2021) Development of integrated sustainability performance indicators for better management of smart cities. *Sustainable Cities and Society* 67: 102704.

Adam T (2021) *What geographic inequalities in COVID-19 mortality rates and health can tell us about levelling up*. Available at: https://www.health.org.uk/news-and-comment/charts-and-infographics/what-geographic-inequalities-in-covid-19-mortality-rates-can-tell-us-about-levelling-up (accessed 09/2021).

Ahmed ME, Rabin MRI and Chowdhury FN (2020a) COVID-19: Social Media Sentiment Analysis on Reopening. *arXiv preprint arXiv:2006.00804*.

Ahmed U, Moreno AT and Moeckel R (2020b) Microscopic activity sequence generation: a multiple correspondence analysis to explain travel behavior based on socio-demographic person attributes. *Transportation*. DOI: 10.1007/s11116-020-10103-1.

Al-Shabi MA (2020) Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions Mining. *International Journal of Computer Science and Network Security* 20(1): 51-57.

Alberti M, McPhearson T, and Gonzalez A (2018) *Embracing urban complexity. Urban planet: knowledge towards sustainable cities.* First edition. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/9781316647554.004, 45-67.

Aldred R and Goodman A (2020) Low Traffic Neighbourhoods, Car Use, and Active Travel: Evidence from the People and Places Survey of Outer London Active Travel Interventions. *Transport Findings*.

Aleta NB, Alonso CM and Ruiz RMA (2017) Smart Mobility and Smart Environment in the Spanish cities. *3rd Conference on Sustainable Urban Mobility (3rd Csum 2016)* 24: 163-170.

Alexander ER (2020) Complexity, institutions and institutional design. *Handbook on Planning and Complexity*. Edward Elgar Publishing.

Allahviranloo M and Aissaoui L (2019) A comparison of time-use behavior in metropolitan areas using pattern recognition techniques. *Transportation Research Part a-Policy and Practice* 129: 271-287.

Allahviranloo M, Regue R and Recker W (2017) Modeling the activity profiles of a population. *Transportmetrica B-Transport Dynamics* 5(4): 426-449.

Allen PM (1997) *Cities and Regions as Self-organizing Systems: Models of Complexity.* Psychology Press.

Anda C, Erath A and Fourie PJ (2017) Transport modelling in the age of big data. *International Journal of Urban Sciences* 21: 19-42.

Anke J, Francke A, Schaefer LM, et al. (2021) Impact of SARS-CoV-2 on the mobility behaviour in Germany. *European Transport Research Review* 13(1).

Anthony B, Petersen SA, Ahlers D, et al. (2020) Big data driven multi-tier architecture for electric mobility as a service in smart cities A design science approach. *International Journal of Energy Sector Management* 14(5): 1023-1047.

Anthopoulos L (2017) Smart utopia VS smart reality: Learning by experience from 10 smart city cases. *Cities* 63: 128-148.

Anthopoulos L, Janssen M and Weerakkody V (2019) A Unified Smart City Model (USCM) for smart city conceptualization and benchmarking. *Smart cities and smart spaces: Concepts, methodologies, tools, and applications*. IGI Global, pp.247-264.

Anthopoulos LG, Pourzolfaghar Z, Lemmer K, et al. (2022) Smart cities as hubs: Connect, collect and control city flows. *Cities* 125: 103660.

Argyriou I and Barry J (2021) The political economy of socio-technical transitions: A relational view of the state and bus system decarbonization in the United Kingdom. *Energy Research & Social Science* 79: 102174.

Aria M and Cuccurullo C (2017) bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics* 11(4): 959-975.

Arun R, Suresh V, Veni Madhavan CE, et al. (2010) On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. Berlin, Heidelberg: Springer Berlin Heidelberg, 391-402.

Audouin M and Finger M (2018) The development of Mobility-as-a-Service in the Helsinki metropolitan area: A multi-level governance analysis. *Research in Transportation Business and Management* 27: 24-35.

Avineri E (2016) Complexity theory and transport planning: Fractal traffic networks. In *A Planner's Encounter with Complexity*. Routledge.

Balduini M, Brambilla M, Della Valle E, et al. (2019) Models and Practices in Urban Data Science at Scale. *Big Data Research* 17: 66-84.

Bari C, Chopade R, Kachwa S, et al. (2021) Impact of COVID-19 on educational trips - an Indian case study. *Transportation Letters-the International Journal of Transportation Research* 13(5-6): 375-387.

Battarra R, Gargiulo C, Tremiterra MR, et al. (2018a) Smart mobility in Italian metropolitan cities: A comparative analysis through indicators and actions. *Sustainable Cities and Society* 41: 556-567.

Battarra R, Pinto F and Tremiterra MR (2018b) Indicators and Actions for the Smart and Sustainable City: A Study on Italian Metropolitan Cities. *Smart Planning: Sustainability and Mobility in the Age of Change*. DOI: 10.1007/978-3-319-77682-8_6. 83-107.

Batty M (2010) Towards a new science of cities. Taylor & Francis.

Batty M (2012) Managing complexity, reworking prediction. *Environment and Planning B-Planning & Design* 39(4): 607-608.

Batty M (2013) Big data, smart cities and city planning. *Dialogues in Human Geography* 3(3): 274-279.

Batty M (2019) Urban analytics defined. *Environment and Planning B-Urban Analytics and City Science* 46(3): 403-405.

Batty M (2020) The Coronavirus crisis: What will the post-pandemic city look like? *Environment and Planning B-Urban Analytics and City Science* 47(4): 547-552.

Batty M, Axhausen KW, Giannotti F, et al. (2012) Smart cities of the future. *The European Physical Journal Special Topics* 214(1): 481-518.

Batty M and Marshall S (2012) The origins of complexity theory in cities and planning. *Complexity theories of cities have come of age*. Springer, pp.21-45.

BBC (2020a) Coronavirus: Capt Tom Moore's NHS fundraiser hits £17m. 17/04/2020.

BBC (2020b) Coronavirus: Five London bus workers die, union confirms. 04/04/2020.

Ben-Akiva M and Abou-Zeid M (2013) Methodological issues in modelling time-of-travel preferences. *Transportmetrica a-Transport Science* 9(9): 846-859.

Benavente-Peces C and Ibadah N (2020) Buildings Energy Efficiency Analysis and Classification Using Various Machine Learning Technique Classifiers. *Energies* 13(13): 3497.

Benevolo C, Dameri RP and D'Auria B (2016) Smart Mobility in Smart City Action Taxonomy, ICT Intensity and Public Benefits. *Empowering Organizations: Enabling Platforms and Artefacts* 11: 13-28.

Bettencourt LMA (2014) The Uses of Big Data in Cities. *Big Data* 2(1): 12-22.

Bhandari N and Pahwa P (2020) Evaluating performance of agglomerative clustering for extended NMF. *Journal of Statistics and Management Systems* 23(7): 1117-1128.

Bibri SE (2018) A foundational framework for smart sustainable city development: Theoretical, disciplinary, and discursive dimensions and their synergies. *Sustainable Cities and Society* 38: 758-794.

Bibri SE (2019) The Sciences Underlying Smart Sustainable Urbanism: Unprecedented Paradigmatic and Scholarly Shifts in Light of Big Data Science and Analytics. *Smart Cities* 2(2): 179-213.

Bibri SE (2020) Compact urbanism and the synergic potential of its integration with data-driven smart urbanism : An extensive interdisciplinary literature review. *Land Use Policy* 97: 104703.

Bibri SE and Krogstie J (2017a) On the social shaping dimensions of smart sustainable cities: A study in science, technology, and society. *Sustainable Cities and Society* 29: 219-246.

Bibri SE and Krogstie J (2017b) Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable Cities and Society* 31: 183-212.

Billones RKC, Guillermo MA, Lucas KC, et al. (2021) Smart Region Mobility Framework. *Sustainability* 13(11).

Blei DM, Ng AY and Jordan MI (2003) Latent dirichlet allocation. *the Journal of machine Learning research* 3: 993-1022.

Blumer H (1948) Public Opinion and Public Opinion Polling. *American Sociological Review* 13(5): 542-549.

Bojovic D, Benavides J and Soret A (2020) What we can learn from birdsong: Mainstreaming teleworking in a post-pandemic world. *Earth System Governance* 5.

Bolin JH, Edwards JM, Finch WH, et al. (2014) Applications of cluster analysis to the creation of perfectionism profiles: a comparison of two clustering approaches. *Frontiers in Psychology* 5.

Bolivar MPR and Meijer AJ (2016) Smart Governance: Using a Literature Review and Empirical Analysis to Build a Research Model. *Social Science Computer Review* 34(6): 673-692.

Brock G, Pihur V, Datta S, et al. (2008) clValid: An R package for cluster validation. *Journal of Statistical Software* 25: 1-22.

Brum-Bastos VS, Long JA and Demšar U (2018) Weather effects on human mobility: a study using multi-channel sequence analysis. *Computers, Environment and Urban Systems* 71: 131-152.

Budd L and Ison S (2020) Responsible Transport: A post-COVID agenda for transport policy and practice. *Transportation Research Interdisciplinary Perspectives* 6: 100151.

Buehler R and Pucher J (2021) COVID-19 Impacts on Cycling, 2019-2020. *Transport Reviews* 41(4): 393-400.

Buhat CAH, Lutero DS, Olave YH, et al. (2020) Modeling the Transmission of Respiratory Infectious Diseases in Mass Transportation Systems. *medRxiv*. DOI: 10.1101/2020.06.09.20126334. 2020.2006.2009.20126334.

Byrne D and Callaghan G (2013) *Complexity theory and the social sciences: The state of the art.* Routledge.

Cairney P (2012) Complexity Theory in Political Science and Public Policy. *Political Studies Review* 10(3): 346-358.

Cambridgeshire and Peterborough Combined Authority (2021) *Cambridgeshire and Peterborough Combined Authority.* Available at: https://cambridgeshirepeterborough-ca.gov.uk/ (Accessed 12/2022)

Campaign for Better Transport (2020) Covid-19 Recovery: Renewing the transport system. Reportno. Report Number|, Date. Place Published|: Institution|.

Cao J, Xia T, Li J, et al. (2009) A density-based method for adaptive LDA model selection. *Neurocomputing* 72(7-9): 1775-1781.

Capela FD and Ramirez-Marquez JE (2019) Detecting urban identity perception via newspaper topic modeling. *Cities* 93: 72-83.

Castillo H and Pitfield DE (2010) ELASTIC - A methodological framework for identifying and selecting sustainable transport indicators. *Transportation Research Part D-Transport and Environment* 15(4): 179-188.

Cellina F, Castri R, Simao JV, et al. (2020) Co-creating app-based policy measures for mobility behavior change: A trigger for novel governance practices at the urban level. *Sustainable Cities and Society* 53.

Centre for Connected and Autonomous Vehicles (2020) Centre for Connected and Autonomous Vehicles.

Chan AKM, Nickson CP, Rudolph JW, et al. (2020) Social media for rapid knowledge dissemination: early experience from the COVID-19 pandemic. *Anaesthesia*. DOI: 10.1111/anae.15057.

Chen R, Zhang J, Ravishanker N, et al. (2019) Clustering activity–travel behavior time series using topological data analysis. *Journal of Big Data Analytics in Transportation* 1(2): 109-121.

Chen Y, Niu H and Silva EA (2022) The Road to Recovery: sensing public opinion towards reopening measures with social media data in post-lockdown cities. *Cities*, 132, 104054.

Chen Y and Silva EA (2021) Smart transport: A comparative analysis using the most used indicators in the literature juxtaposed with interventions in English metropolitan areas. *Transportation Research Interdisciplinary Perspectives* 10: 100371.

Chen Y and Silva EA (forthcoming) How can Complexity Theory and Data Science assist Smart City Governance? A review.

Chen Y, Silva EA and Reis JP (2020) Measuring policy debate in a regrowing city by sentiment analysis using online media data: a case study of Leipzig 2030. *Regional Science Policy & Practice*.

Cheng, I., Heyl, J., Lad, N., Facini, G., & Grout, Z. (2021). Evaluation of Twitter data for an emerging crisis: an application to the first wave of COVID-19 in the UK. *Scientific Reports*, 11(1), 1-13.

Cheng L, Chen XW, Yang S, et al. (2019) Structural equation models to analyze activity participation, trip generation, and mode choice of low-income commuters. *Transportation Letters-the International Journal of Transportation Research* 11(6): 341-349.

Cho SH and Park HC (2021) Exploring the Behaviour Change of Crowding Impedance on Public Transit due to COVID-19 Pandemic: Before and After Comparison. *Transportation Letters-the International Journal of Transportation Research* 13(5-6): 367-374.

Cho SJ, Janssens D, Joh CH, et al. (2019) Space-Time Sequential Similarity for Identifying Factors of Activity-Travel Pattern Segmentation: A Measure of Sequence Alignment and Path Similarity. *Geographical Analysis* 51(2): 203-220.

Cinelli M, Quattrociocchi W, Galeazzi A, et al. (2020) The covid-19 social media infodemic. *arXiv preprint arXiv:2003.05004*.

Cleveland WS (2014) Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics. *Statistical Analysis and Data Mining* 7(6): 414-417.

Colding J, Colding M and Barthel S (2020) The smart city model: A new panacea for urban sustainability or unmanageable complexity? *Environment and Planning B-Urban Analytics and City Science* 47(1): 179-187.

Collins JP (2010) The Fourth Paradigm Data-Intensive Scientific Discovery. *Science* 327(5972): 1455-1456.

Colneric N and Demsar J (2020) Emotion Recognition on Twitter: Comparative Study and Training a Unison Model. *IEEE transactions on affective computing* 11(3): 433-446.

Cottrill C, Gault P, Yeboah G, et al. (2017) Tweeting Transit: An examination of social media strategies for transport information management during a large event. *Transportation Research Part C-Emerging Technologies* 77: 421-432.

Crainic TG, Perboli G, Rosano M, et al. (2019) *Transportation for Smart Cities: A Systematic Review.* CIRRELT Montreal, Canada.

Crawford F, Watling DP and Connors RD (2018) Identifying road user classes based on repeated trip behaviour using Bluetooth data. *Transportation Research Part a-Policy and Practice* 113: 55-74.

Cruz R, Jardim J, Mira J, et al. (2018) Smart Rail for Smart Mobility. IEEE, 1-7.

Daisy NS (2018) Microsimulation of activity participation, tour complexity, and mode choice within an activity-based travel demand model system.

Das S, Boruah A, Banerjee A, et al. (2021) Impact of COVID-19: A radical modal shift from public to private transport mode. *Transport Policy* 109: 1-11.

Dave RN (1991) Characterization and Detection of Noise in Clustering. *Pattern Recognition Letters* 12(11): 657-664.

De Haas M, Faber R and Hamersma M (2020) How COVID-19 and the Dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. *Transportation Research Interdisciplinary Perspectives* 6: 100150.

De Roo G (2016) Being or becoming? That is the question! Confronting complexity with contemporary planning theory. *A planner's encounter with complexity*. Routledge, pp.19-40.

De Roo G (2020) Introduction to the Handbook on Planning and Complexity. *Handbook on Planning and Complexity*. Edward Elgar Publishing.

De Roo G and Silva EA (2010) *A planner's encounter with complexity.* Ashgate Publishing, Ltd.

Debnath AK, Chin HC, Haque MM, et al. (2014) A methodological framework for benchmarking smart transport cities. *Cities* 37: 47-56.

Dennis S, Garrett P, Yim H, et al. (2019) Privacy versus open science. *Behavior Research Methods* 51(4): 1839-1848.

Department for Business EaIS (2020) Statistics at BEIS.

Department for Transport (2017) Intelligent Transport Systems in the UK (Progress Report). London, 153.

Department for Transport (2020a) National Travel Survey, 2002-2019: Special Licence Access, [data collection], UK Data Service, 9th Edition. In: Department for Transport (ed).

Department for Transport (2020b) Statistics at DfT.

Department for Transport (2020c) Transport Secretary announces new measures to "keep passengers safe now and level up for the future". GOV.UK.

Deveaud R, SanJuan E and Bellot P (2014) Accurate and effective latent concept modeling for ad hoc information retrieval. *Document numérique* 17(1): 61-84.

Dhar V (2013) Data Science and Prediction. *Communications of the Acm* 56(12): 64-73.

Dharmowijoyo DB, Susilo YO and Karlström A (2016) Day-to-day variability in travellers' activity-travel patterns in the Jakarta metropolitan area. *Transportation* 43(4): 601-621.

Dharmowijoyo DBE, Susilo YO and Karlstrom A (2017) Analysing the complexity of day-to-day individual activity-travel patterns using a multidimensional sequence alignment model: A case study in the Bandung Metropolitan Area, Indonesia. *Journal of Transport Geography* 64: 1-12.

Dingil AE and Esztergar-Kiss D (2021) The Influence of the Covid-19 Pandemic on Mobility Patterns: The First Wave's Results. *Transportation Letters-the International Journal of Transportation Research* 13(5-6): 434-446.

Docherty I (2018) New governance challenges in the era of 'Smart'mobility. *Governance of the smart mobility transition*. 19-32.

Docherty I, Marsden G and Anable J (2018) The governance of smart mobility. *Transportation Research Part A: Policy and Practice* 115: 114-125.

Donoho D (2017) 50 Years of Data Science. *Journal of Computational and Graphical Statistics* 26(4): 745-766.

Drummond P (2021) Assessing City Governance for Low-Carbon Mobility in London. *Sustainability* 13(5).

Dumbliauskas V and Grigonis V (2020) An Empirical Activity Sequence Approach for Travel Behavior Analysis in Vilnius City. *Sustainability* 12(2).

Eisenberg-Guyot J, Peckham T, Andrea SB, et al. (2020) Life-course trajectories of employment quality and health in the US: A multichannel sequence analysis. *Social Science & Medicine* 264: 113327.

Ekman U (2018) Smart City Planning: Complexity. *International Journal of E-Planning Research* 7(3): 1-21.

Elzinga CH and Liefbroer AC (2007) De-standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis. *European Journal of Population / Revue européenne de Démographie* 23(3): 225-250.

Entwisle B and Elias P (2013) *New data for understanding the human condition: International perspectives*.

Eppel EA and Rhodes ML (2018) Complexity theory and public management: a 'becoming' field. *Public Management Review* 20(7): 949-959.

Ezugwu AE, Shukla AK, Agbaje MB, et al. (2021) Automatic clustering algorithms: a systematic review and bibliometric analysis of relevant literature. *Neural Computing & Applications* 33(11): 6247-6306.

Faber A, Rehm SV, Hernandez-Mendez A, et al. (2018) Modeling and Visualizing Smart City Mobility Business Ecosystems: Insights from a Case Study. *Information* 9(11).

Fairnie GA, Wilby DJR and Saunders LE (2016) Active travel in London: The role of travel survey data in describing population physical activity. *Journal of Transport & Health* 3(2): 161-172.

Farooq A, Xie MW, Stoilova S, et al. (2019) The Application of Smart Urban Mobility Strategies and Initiatives: Application to Beijing. *European Transport-Trasporti Europei*.(71).

Feezell JT (2018) Agenda Setting through Social Media: The Importance of Incidental News Exposure and Social Filtering in the Digital Era. *Political Research Quarterly* 71(2): 482-494.

Feng DD, Cheng L and Du MY (2020) Exploring the Impact of Dockless Bikeshare on Docked Bikeshare-A Case Study in London. *Sustainability* 12(15).

Fenwick J and Johnston L (2020) Leading the combined authorities in England: a new future for elected mayors? *Public Money & Management* 40(1): 14-20.

Fernandez-Anez V, Fernández-Güell JM and Giffinger R (2018) Smart City implementation and discourses: An integrated conceptual model. The case of Vienna. *Cities* 78: 4-16.

Ferraro MB and Giordani P (2020) Soft clustering. *Wiley Interdisciplinary Reviews-Computational Statistics* 12(1).

Ferraro MB, Giordani P and Serafini A (2019) fclust: An R Package for Fuzzy Clustering. *R J.* 11(1): 198.

Field C and Jon I (2021) E-Scooters: A New Smart Mobility Option? The Case of Brisbane, Australia. *Planning Theory & Practice* 22(3): 368-396.

Finger M and Audouin M (2019) The Governance of Smart Transportation Systems. Springer.

Folke C (2006) Resilience: The emergence of a perspective for social-ecological systems analyses. *Global Environmental Change-Human and Policy Dimensions* 16(3): 253-267.

Fonzone A, Saleh W and Rye T (2018) Smart urban mobility - Escaping the technological Sirens. *Transportation Research Part a-Policy and Practice* 115: 1-3.

Friis F (2020) An alternative explanation of the persistent low EV-uptake: The need for interventions in current norms of mobility demand. *Journal of Transport Geography* 83.

Gabadinho A, Ritschard G, Mueller NS, et al. (2011) Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software* 40(4): 1-37.

Gabadinho A, Ritschard G, Studer M, et al. (2009) Mining sequence data in R with the TraMineR package: A user’s guide. *Geneva: Department of Econometrics and Laboratory of Demography, University of Geneva*.

Gao J and O’Neill BC (2020) Mapping global urban land for the 21st century with data-driven simulations and Shared Socioeconomic Pathways. *Nature Communications* 11(1): 2302.

Garau C, Desogus G and Coni M (2019) Fostering and Planning a Smart Governance Strategy for Evaluating the Urban Polarities of the Sardinian Island (Italy). *Sustainability* 11(18).

Garau C, Masala F and Pinna F (2015) Benchmarking Smart Urban Mobility: A Study on Italian Cities. *Computational Science and Its Applications - Iccsa 2015, Pt Ii* 9156: 612-623.

Garau C, Masala F and Pinna F (2016) Cagliari and smart urban mobility: Analysis and comparison. *Cities* 56: 35-46.

Geels FW (2005) *Technological transitions and system innovations: a co-evolutionary and socio-technical analysis.* Edward Elgar Publishing.

Geels FW (2012) A socio-technical analysis of low-carbon transitions: introducing the multi-level perspective into transport studies. *Journal of Transport Geography* 24: 471-482.

Geels FW (2020) Micro-foundations of the multi-level perspective on socio-technical transitions: Developing a multi-dimensional model of agency through crossovers between social constructivism, evolutionary economics and neo-institutional theory. *Technological Forecasting and Social Change* 152: 119894.

Geyer R and Cairney P (2015) *Handbook on complexity and public policy.* Edward Elgar Publishing.

Ghosh A, Nundy S, Ghosh S, et al. (2020) Study of COVID-19 pandemic in London (UK) from urban context. *Cities* 106: 102928.

Giffinger R, Fertner C, Kramar H, et al. (2007) City-ranking of European medium-sized cities. *Cent. Reg. Sci. Vienna UT*. 1-12.

Girones ES, van Est R and Verbong G (2019) Transforming mobility: The Dutch smart mobility policy as an example of a transformative STI policy. *Science and Public Policy* 46(6): 820-833.

Gkiotsalitis K and Cats O (2021) Public transport planning adaption under the COVID-19 pandemic crisis: literature review of research needs and directions. *Transport Reviews* 41(3): 374-392.

Gkiotsalitis K and Stathopoulos A (2020) Predicting Traveling Distances and Unveiling Mobility and Activity Patterns of Individuals from Multisource Data. *Journal of Transportation Engineering Part a-Systems* 146(5).

GLA Opinion Research (2020a) COVID-19 online diary - weeks 1 and 2

GLA Opinion Research (2020b) COVID-19 online diary - weeks 3 and 4

GLA Opinion Research (2020c) COVID-19 online diary - weeks 5 and 6

GLA Opinion Research (2020d) COVID-19 online diary - weeks 7 and 8

GLA Opinion Research (2020e) COVID-19 online diary: Summary report. London Datastore.

Gong K, Zhang L, Ni D, et al. (2020) An expert system to discover key congestion points for urban traffic. *Expert Systems with Applications* 158: 113544.

Gonzalez RA, Ferro RE and Liberona D (2020) Government and governance in intelligent cities, smart transportation study case in Bogota Colombia. *Ain Shams Engineering Journal* 11(1): 25-34.

Google (2020) *COVID-19 community mobility reports*.

Gottlieb M and Dyer S (2020) Information and Disinformation: Social Media in the COVID-19 Crisis. *Academic Emergency Medicine* 27(7): 640-641.

Goulias KG (2021) Special issue on understanding the relationships between COVID-19 and transportation. *Transportation Letters-the International Journal of Transportation Research* 13(5-6): 327-330.

Government Office for Science (2019) *A time of unprecedented change in the transport system: the Future of Mobility*.

Grant-Muller SM, Gal-Tzur A, Minkov E, et al. (2015) Enhancing transport data collection through social media sources: methods, challenges and opportunities for textual data. *Iet Intelligent Transport Systems* 9(4): 407-417.

Greater London Authority (2020) Coronavirus (COVID-19) Mobility Report. London Datastore.

Greater London Authority and London Office of Technology & Innovation (2020) Written evidence from the Greater London Authority (GLA) and the London Office of Technology & Innovation (LOTI) at London Councils (DTA 24). In: Committee PAaCA (ed).

Greater Manchester Combined Authority (2022) *Mayor accelerates London-style transport revolution; bus journeys capped at £1 for kids, £2 adults*. Available at: https://www.greatermanchester-ca.gov.uk/news/mayor-accelerates-london-style-transport-revolution-bus-journeys-capped-at-1-for-kids-2-adults/.

Greater Manchester Combined Authority (2020) *Greater Manchester Digital Blueprint*. Available at: https://www.greatermanchester-ca.gov.uk/media/6275/gmca_blueprint_jul-22.pdf (Accessed 12/2022)

Griffiths S, Del Rio DF and Sovacool B (2021) Policy mixes to achieve sustainable mobility after the COVID-19 crisis. *Renewable & Sustainable Energy Reviews* 143.

Griffiths TL and Steyvers M (2004) Finding scientific topics. *Proceedings of the National academy of Sciences* 101(suppl 1): 5228-5235.

Grun B and Hornik K (2011) Topicmodels: An R Package for Fitting Topic Models. *Journal of Statistical Software* 40(13): 1-30.

Guo Y, Tang Z and Guo J (2020) Could a Smart City Ameliorate Urban Traffic Congestion? A Quasi-Natural Experiment Based on a Smart City Pilot Program in China. *Sustainability* 12(6): 2291.

Habib KN (2018) A comprehensive utility-based system of activity-travel scheduling options modelling (CUSTOM) for worker's daily activity scheduling processes. *Transportmetrica a-Transport Science* 14(4): 292-315.

Habib KN, Hawkins J, Shakib S, et al. (2021) Assessing the impacts of COVID-19 on urban passenger travel demand in the greater Toronto area: description of a multi-pronged and multi-staged study with initial results. *Transportation Letters-the International Journal of Transportation Research* 13(5-6): 353-366.

Habitat U (2020) World Cities Report 2020: The value of sustainable urbanization. *Nairobi, Kenya*.

Hadjidemetriou GM, Sasidharan M, Kouyialis G, et al. (2020) The impact of government measures and human mobility trend on COVID-19 related deaths in the UK. *Transportation Research Interdisciplinary Perspectives* 6: 100167.

Hafezi MH, Daisy NS, Liu L, et al. (2018a) Daily activity and travel sequences of students, faculty and staff at a large Canadian university. *Transportation Planning and Technology* 41(5): 536-556.

Hafezi MH, Daisy NS, Millward H, et al. (2021) Framework for development of the Scheduler for Activities, Locations, and Travel (SALT) model. *Transportmetrica a-Transport Science*. DOI: 10.1080/23249935.2021.1921879.

Hafezi MH, Liu L and Millward H (2017) Identification of Representative Patterns of Time Use Activity Through Fuzzy C-Means Clustering. *Transportation Research Record*. DOI: 10.3141/2668-05.(2668): 38-50.

Hafezi MH, Liu L and Millward H (2018b) Learning Daily Activity Sequences of Population Groups using Random Forest Theory. *Transportation Research Record* 2672(47): 194-207.

Hafezi MH, Liu L and Millward H (2019) A time-use activity-pattern recognition model for activity-based travel demand modeling. *Transportation* 46(4): 1369-1394.

Hajek P, Youssef A and Hajkova V (2022) Recent developments in smart city assessment: A bibliometric and content analysis-based literature review. *Cities*. DOI: https://doi.org/10.1016/j.cities.2022.103709. 103709.

Haken H, Portugali J, Batty M, et al. (2021) Epilogue: cities and complexity in the time of COVID-19. *Handbook on Cities and Complexity*. 391.

Hall P (2009) Looking Backward, Looking Forward: The City Region of the Mid-21st Century. *Regional Studies* 43(6): 803-817.

Harriss L and Kearney P (2021) Smart Cities. In: Parliament U (ed). POST.

Heink U and Kowarik I (2010) What are indicators? On the definition of indicators in ecology and environmental planning. *Ecological indicators*, 10(3), 584-593.

Hey AJ, Tansley S and Tolle KM (2009) *The fourth paradigm: data-intensive scientific discovery.* Microsoft research Redmond, WA.

Hickman H and While A (2017) Combined authorities: Signs of success.

Highways England (2020) Smart motorways.

HM Government (2018) North of Tyne Combined Authority devolution deal.

Hochtl J, Parycek P and Schollhammer R (2016) Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce* 26(1-2): 147-169.

Hollander JB and Renski H (2017) Measuring urban attitudes embedded in microblogging data: shrinking versus growing cities. *Town Planning Review* 88(4): 465-490.

Hong A, Baker L, Curiel RP, et al. (2022) Reconciling big data and thick data to advance the new urban science and smart city governance. *Journal of Urban Affairs*. DOI: 10.1080/07352166.2021.2021085.

Hussain AA, Bouachir O, Al-Turjman F, et al. (2020) AI techniques for COVID-19. *Ieee Access* 8: 128776-128795.

Icasiano CDA and Taeihagh A (2021) Governance of the Risks of Ridesharing in Southeast Asia: An In-Depth Analysis. *Sustainability* 13(11).

Innes JE and Booher DE (2010) *Planning with complexity: An introduction to collaborative rationality for public policy.* Routledge.

Ishwarappa and Anuradha J (2014) A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology. *1st International Conference on Intelligent Computing, Communication and Convergence (ICCC).* Bhubaneshwar, INDIA, 319-324.

Isoaho K, Gritsenko D and Makela E (2021) Topic Modeling and Text Analysis for Qualitative Policy Research. *Policy Studies Journal* 49(1): 300-324.

Jacobs J (1961) *The Death and Life of Great American Cities.* Random House, New York.

Jiang H (2021) Smart urban governance in the ‘smart’era: Why is it urgently needed?. *Cities*, 111, 103004.

Jiang S, Ferreira J and Gonzalez MC (2012) Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery* 25(3): 478-510.

Jiao JF and Azimian A (2021) Exploring the factors affecting travel behaviors during the second phase of the COVID-19 pandemic in the United States. *Transportation Letters-the International Journal of Transportation Research* 13(5-6): 331-343.

Joyce P (2021) Public governance, agility and pandemics: a case study of the UK response to COVID-19. *International Review of Administrative Sciences* 87(3): 536-555.

Kaaristo M, Medway D, Burton J, et al. (2020) Governing mobilities on the UK canal network. *Mobilities* 15(6): 844-861.

Kanda W and Kivimaa P (2020) What opportunities could the COVID-19 outbreak offer for sustainability transitions research on electricity and mobility? *Energy Research & Social Science* 68: 101666.

Kandt J and Batty M (2021) Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities* 109: 102992.

Kang W, Oshan T, Wolf LJ, et al. (2019) A roundtable discussion: Defining urban data science. *Environment and Planning B-Urban Analytics and City Science* 46(9): 1756-1768.

Katakis I (2015) Mining urban data (part A) Preface. *Information Systems* 54: 113-114.

Kato S and Ahern J (2008) 'Learning by doing': adaptive planning as a strategy to address uncertainty in planning. *Journal of Environmental Planning and Management* 51(4): 543-559.

Kaufman L and Rousseeuw PJ (2009) *Finding groups in data: an introduction to cluster analysis.* John Wiley & Sons.

Kearney MW (2020) rtweet: Collecting Twitter data. *R package version 0.7.0* 7.

Kester J (2018) Governing electric vehicles: mobilizing electricity to secure automobility. *Mobilities* 13(2): 200-215.

Kim J (2022) Smart city trends: A focus on 5 countries and 15 companies. *Cities* 123: 103551.

Kim K (2014) Discrepancy Analysis of Activity Sequences What Explains the Complexity of People's Daily Activity-Travel Patterns? *Transportation Research Record*. DOI: 10.3141/2413-03.(2413): 24-33.

Kim K (2018) Recent Advances in Activity-Based Travel Demand Models for Greater Flexibility.

Kitchin R (2016) The ethics of smart cities and urban science. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences* 374(2083).

Kitchin R (2019) The Timescape of Smart Cities. *Annals of the American Association of Geographers* 109(3): 775-790.

Kitchin R, Lauriault TP and McArdle G (2015) Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies Regional Science* 2(1): 6-28.

Kolarova V, Eisenmann C, Nobis C, et al. (2021) Analysing the impact of the COVID-19 outbreak on everyday travel behaviour in Germany and potential implications for future travel patterns. *European Transport Research Review* 13(1).

Kontokosta CE (2021) Urban informatics in the science and practice of planning. *Journal of Planning Education and Research* 41(4): 382-395.

Koushik ANP, Manoj M and Nezamuddin N (2020) Machine learning applications in activity-travel behaviour research: a review. *Transport Reviews* 40(3): 288-311.

Kumar H, Singh MK and Gupta MP (2018) Smart mobility: Crowdsourcing solutions for smart transport system in smart cities context. 482-488.

Laker L (2020) World Cities Turn Their Streets over to Walkers and Cyclists. *The Guardian*, 11 April 2020.

Landrigan PJ, Fuller R, Hu H, et al. (2018) Pollution and Global Health - An Agenda for Prevention. *Environmental Health Perspectives* 126(8).

Langlois GG, Koutsopoulos HN and Zhao JH (2016) Inferring patterns in the multi-week activity sequences of public transport users. *Transportation Research Part C-Emerging Technologies* 64: 1-16.

Laverty AA, Millett C, Majeed A, et al. (2020) COVID-19 presents opportunities and threats to transport and health. *Journal of the Royal Society of Medicine* 113(7): 251-254.

Lawrenz F, Kollmann EK, King JA, et al. (2018) Promoting evaluation capacity building in a complex adaptive system. *Evaluation and Program Planning*, 69, 53-60.

Li H, Zhang Y, Zhu M, et al. (2021) Impacts of COVID-19 on the usage of public bicycle share in London. *Transportation Research Part A: Policy and Practice* 150: 140-155.

Li JM and Lewis HW (2016) Fuzzy Clustering Algorithms - Review of the Applications. *2016 Ieee International Conference on Smart Cloud (Smartcloud)*. DOI: 10.1109/SmartCloud.2016.14. 282-288.

Li WW, Batty M and Goodchild MF (2019a) Real-time GIS for smart cities. *International Journal of Geographical Information Science*. DOI: 10.1080/13658816.2019.1673397.

Li X, Fong PSW, Dai SL, et al. (2019b) Towards sustainable smart cities: An empirical comparative assessment and development pattern optimization in China. *Journal of Cleaner Production* 215: 730-743.

Lin S, Hsiao YY, and Wang M (2014) *Test review: the profile of mood states 2nd edition*.

Liu L, Silva EA and Yang ZS (2021) Similar outcomes, different paths: Tracing the relationship between neighborhood-scale built environment and travel behavior using activity-based modelling. *Cities* 110.

Liu Y and Cheng T (2017) Characterising Passengers' Travel Patterns in London Public Transit. GISRUK.

Liverpool City Region (2021) *LCR Digital Strategy & Action Plan.* Available at: https://www.liverpoolcityregion-ca.gov.uk/wp-content/uploads/LCR-Digital-Strategy-05-2.pdf (Accessed 12/2022)

Loorbach D, Schwanen T, Doody BJ, et al. (2021) Transition governance for just, sustainable urban mobility: An experimental approach from Rotterdam, the Netherlands. *Journal of Urban Mobility* 1: 100009.

Loper E and Bird S (2002) NLTK: the natural language toolkit. *arXiv preprint cs/0205028*.

Lopez-Carreiro I and Monzon A (2018) Evaluating sustainability and innovation of mobility patterns in Spanish cities. Analysis by size and urban typology. *Sustainable Cities and Society* 38: 684-696.

Lorencka M and Obrebska M (2018) English Combined Authorities and Italian Metropolitan Cities: A Comparative Perspective. *Romanian Journal of Political Science* 18(2): 119-148.

Lovelace R, Morgan M, Talbot J, et al. (2020) Methods to prioritise pop-up active transport infrastructure and their application in a national cycleway prioritisation tool.

Lu Y and Zhang LL (2020) Social media WeChat infers the development trend of COVID-19. *Journal of Infection* 81(1): E82-E83.

Lyons G (2018) Getting smart about urban mobility - Aligning the paradigms of smart and sustainable. *Transportation Research Part a-Policy and Practice* 115: 4-14.

Lyons G and Davidson C (2016) Guidance for transport planning and policymaking in the face of an uncertain future. *Transportation Research Part a-Policy and Practice* 88: 104-116.

Ma Y, Lan J, Thornton T, et al. (2018) Challenges of collaborative governance in the sharing economy: The case of free-floating bike sharing in Shanghai. *Journal of Cleaner Production* 197: 356-365.

Mackintosh T (2020) Why Transport for London's finances are far from healthy. *BBC*.

Maechler M, Rousseeuw P, Struyf A, et al. (2013) Package 'cluster'. *Dosegljivo na*.

Maier MJ (2014) DirichletReg: Dirichlet regression for compositional data in R.

Manders TNT, Wieczorek AJA and Verbong GPJG (2020) Complexity, tensions, and ambiguity of intermediation in a transition context: The case of Connecting Mobility. *Environmental Innovation and Societal Transitions* 34: 183-208.

Mao YX (2020) Combating COVID-19 Through Collaborative Governance: Lessons from East Asia. *Chinese Public Administration Review* 11(2): 132-141.

Marsden G and Docherty I (2019) *Governance of UK Transport Infrastructures (Future of Mobility: Evidence Review)*.

Marsden G and Docherty I (2021) Mega-disruptions and policy change: Lessons from the mobility sector in response to the Covid-19 pandemic in the UK. *Transport Policy* 110: 86-97.

McBride EC, Davis AW and Goulias KG (2020) Exploration of Statewide Fragmentation of Activity and Travel and a Taxonomy of Daily Time Use Patterns using Sequence Analysis in California. *Transportation Research Record* 2674(12): 38-51.

Meijer A and Bolivar MPR (2016) Governing the smart city: a review of the literature on smart urban governance. *International Review of Administrative Sciences* 82(2): 392-408.

Mellon J and Prosser C (2017) Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics* 4(3): 2053168017720008.

Miller EJ and Roorda MJ (2003) Prototype model of household activity-travel scheduling. *Transportation Research Record* 1831(1): 114-121.

Millward H, Hafezi MH and Daisy NS (2019) Activity travel of population segments grouped by daily time-use: GPS tracking in Halifax, Canada. *Travel Behaviour and Society* 16: 161-170.

Milne D and Watling D (2019) Big data and understanding change in the context of planning transport systems. *Journal of Transport Geography* 76: 235-244.

Mladenovic MN and Haavisto N (2021) Interpretative flexibility and conflicts in the emergence of Mobility as a Service: Finnish public sector actor perspectives. *Case Studies on Transport Policy* 9(2): 851-859.

Mohammadian HD and Rezaie F (2020) Blue-Green Smart Mobility Technologies as Readiness for Facing Tomorrow's Urban Shock toward the World as a Better

Place for Living (Case Studies: Songdo and Copenhagen). *Technologies* 8(3).

Mohanty SP, Choppali U and Kougianos E (2016) Everything you wanted to know about smart cities: The internet of things is the backbone. *IEEE Consumer Electronics Magazine* 5(3): 60-70.

Moiseeva A, Timmermans H, Choi J, et al. (2014) Sequence alignment analysis of variability in activity travel patterns through 8 weeks of diary data. *Transportation Research Record* 2412(1): 49-56.

Moscholidou I and Pangbourne K (2019) A preliminary assessment of regulatory efforts to steer smart mobility in London and Seattle. *Transport Policy*.

Moslem S, Campisi T, Szmelter-Jarosz A, et al. (2020) Best-Worst Method for Modelling Mobility Choice after COVID-19: Evidence from Italy. *Sustainability* 12(17): 19.

Mounce R, Beecroft M and Nelson JD (2020) On the role of frameworks and smart mobility in addressing the rural mobility problem. *Research in Transportation Economics* 83.

Mouratidis K and Papagiannakis A (2021) COVID-19, internet, and mobility: The rise of telework, telehealth, e-learning, and e-shopping. *Sustainable Cities and Society* 74: 103182.

Mouratidis K, Peters S and van Wee B (2021) Transportation technologies, sharing economy, and teleactivities: Implications for built environment and travel. *Transportation Research Part D: Transport and Environment* 92: 102716.

Musselwhite C, Avineri E and Susilo Y (2021) Restrictions on mobility due to the coronavirus Covid19: Threats and opportunities for transport and health. *Journal of Transport & Health* 20.

Nagesh A (2020) The Uber driver evicted from home and left to die of coronavirus. *BBC*, 28/04/2020.

Narlikar A and Sottilotta CE (2021) Pandemic narratives and policy responses: west European governments and COVID-19. *Journal of European Public Policy*. DOI: 10.1080/13501763.2021.1942152.

Nasser N, Khan N, Karim L, et al. (2021) An efficient Time-sensitive data scheduling approach for Wireless Sensor Networks in smart cities. *Computer Communications* 175: 112-122.

National Audit Office (2017) Progress in setting up combined authorities. House of Commons London.

National Health Service (2020) Data, insights and statistics.

National Health Service England (2020a) *Ambulance Global Digital Exemplars.* Available at: https://www.england.nhs.uk/digitaltechnology/connecteddigitalsystems/exemplars/ambulance-global-digital-exemplars/ (accessed 06/2020).

National Health Service England (2020b) *Global Digital Exemplars*. Available at: https://www.england.nhs.uk/digitaltechnology/connecteddigitalsystems/exemplars/ (accessed 06/2020).

Nielsen F (2011) AFINN: A new word list for sentiment analysis on Twitter.

Nikita M, Chaney N and Chaney MN (2020) Package 'ldatuning'.

Nikitas A, Michalakopoulou K, Njoya ET, et al. (2020) Artificial Intelligence, Transport and the Smart City: Definitions and Dimensions of a New Mobility Era. *Sustainability* 12(7).

Nikolaeva A, te Brommelstroet M, Raven R, et al. (2019) Smart cycling futures: Charting a new terrain and moving towards a research agenda. *Journal of Transport Geography* 79.

Nikolaidou A and Papaioannou P (2018) Utilizing social media in transport planning and public transit quality: Survey of literature. *Journal of Transportation Engineering, Part A: Systems* 144(4): 04018007.

Niu HF and Silva EA (2020) Crowdsourced Data Mining for Urban Activity: Review of Data Sources, Applications, and Methods. *Journal of Urban Planning and Development* 146(2).

North East Combined Authority (2021) *What NECA does*. Available at: https://northeastca.gov.uk/what-we-do/ (Accessed 12/2022)

North East Combined Authority. (2022) *The North East Joint Transport Committee.* Available at: https://northeastca.gov.uk/decision-making/the-north-east-joint-transport-committee/ (accessed 12/2022)

North of Tyne (2021) *North of Tyne Economic Vision*. Available at: https://www.northoftyne-ca.gov.uk/wp-content/uploads/2020/09/NorthofTyne_EconomicVision_webfinal.pdf (Accessed 12/2022)

Office for National Statistics (2020) Health and social care.

Office for National Statistics (2021) Deaths due to COVID-19 by local area and deprivation. In: ONS (ed).

Oldbury K and Isaksson K (2021) Governance arrangements shaping driverless shuttles in public transport: The case of Barkarbystaden, Stockholm. *Cities* 113.

Pangbourne K, Mladenovic MN, Stead D, et al. (2020) Questioning mobility as a service: Unanticipated implications for society and governance. *Transportation Research Part a-Policy and Practice* 131: 35-49.

Park D, Jeon YH, Cho SJ, et al. (2020) Profiling of Clusters of Activity-Travel Sequences Using a Genetic Algorithm. *Geographical Analysis*. DOI: 10.1111/gean.12248. 22.

Park W, Choi K and Joh C-H (2018) How does work activity affect quality of life?: A spatial analysis of trip chain behavior. *KSCE Journal of Civil Engineering* 22(1): 320-329.

Pase F, Chiariotti F, Zanella A, et al. (2020) Bike Sharing and Urban Mobility in a Post-Pandemic World. *Ieee Access* 8: 187291-187306.

Pasichnyi O, Levihn F, Shahrokni H, et al. (2019) Data-driven strategic planning of building energy retrofitting: The case of Stockholm. *Journal of Cleaner Production* 233: 546-560.

Paton G (2020) Coronavirus: Tube trains sprayed but services still overcrowded. *The Times*, 27/03/2020.

Petrowski K, Albani C, Zenger M, et al. (2021). Revised short screening version of the Profile of Mood States (POMS) from the German general population. *Frontiers in psychology*, 12, 631668.

Perveen S, Yigitcanlar T, Kamruzzaman M, et al. (2020) How can transport impacts of urban growth be modelled? An approach to consider spatial and temporal scales. *Sustainable Cities and Society* 55: 102031.

Pindarwati A and Wijayanto AW (2015) *Measuring Performance Level of Smart Transportation System in Big Cities of Indonesia Comparative Study: Jakarta, Bandung, Medan, Surabaya, and Makassar.* New York: Ieee.

Pinna F, Masala F and Garau C (2017a) Urban policies and mobility trends in Italian smart cities. *Sustainability* 9(4): 494.

Pinna F, Masala F and Garau C (2017b) Urban Policies and Mobility Trends in Italian Smart Cities. *Sustainability* 9(4).

Politis I, Georgiadis G, Kopsacheilis A, et al. (2021) Capturing Twitter Negativity Pre- vs. Mid-COVID-19 Pandemic: An LDA Application on London Public Transport System. *Sustainability* 13(23): 13356.

Pop MD and Prostean O (2019) Identification of significant metrics and indicators for smart mobility. In: Lemle LD (ed) *International Conference on Applied Sciences*. Bristol: Iop Publishing Ltd.

Portugali J (2012) Complexity Theories of Cities: First, Second or Third Culture of Planning? *Complexity and Planning: Systems, Assemblages and Simulations*. 117-140.

Portugali J, Meyer H, Stolk E, et al. (2012) *Complexity theories of cities have come of age: an overview with implications to urban planning and design.* Springer Science & Business Media.

Prieto M, Baltas G and Stan V (2017) Car sharing adoption intention in urban areas: What are the key sociodemographic drivers? *Transportation Research Part a-Policy and Practice* 101: 218-227.

Prigogine I (1978) Time, structure, and fluctuations. *Science* 201(4358): 777-785.

Public Health England (2020) PHE data and analysis tools.

Puri N, Coomes EA, Haghbayan H, et al. (2020) Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. *Human Vaccines & Immunotherapeutics* 16(11): 2586-2593.

Rauws W (2017) Embracing Uncertainty Without Abandoning Planning Exploring an Adaptive Planning Approach for Guiding Urban Transformations. *Disp* 53(1): 32-45.

Raux C, Ma T-Y and Cornelis E (2016) Variability in daily activity-travel patterns: the case of a one-week travel diary. *European Transport Research Review* 8(4): 1-14.

Resch B and Szell M (2019) Human-Centric Data Science for Urban Studies. *Isprs International Journal of Geo-Information* 8(12).

Ribeiro P, Dias G and Pereira P (2021) Transport Systems and Mobility for Smart Cities. *Applied System Innovation* 4(3).

Ritschard G Measuring the Nature of Individual Sequences. *Sociological Methods & Research* 0(0): 00491241211036156.

Ruhlandt RWS (2018) The governance of smart cities: A systematic literature review. *Cities* 81: 1-23.

Ruspini EH, Bezdek JC and Keller JM (2019) Fuzzy Clustering: A Historical Perspective. *Ieee Computational Intelligence Magazine* 14(1): 45-55.

Saadi I, Mustafa A, Teller J, et al. (2016) Forecasting travel behavior using Markov Chains-based approaches. *Transportation Research Part C-Emerging Technologies* 69: 402-417.

Saberi M, Mahmassani HS, Brockmann D, et al. (2017) A complex network perspective for characterizing urban travel demand patterns: graph theoretical analysis of large-scale origin-destination demand networks. *Transportation* 44(6): 1383-1402.

Sagaris L (2014) Citizen participation for sustainable transport: the case of “Living City” in Santiago, Chile (1997–2012). *Journal of Transport Geography* 41: 74-83.

Samuel J, Rahman MM, Ali GMN, et al. (2020) Feeling Positive About Reopening? New Normal Scenarios from COVID-19 Reopen Sentiment Analytics. *medRxiv*.

Sandford M (2018) The Greater London Authority. *Briefing papers. House of Commons Library, London*.

Sandford M (2019a) Combined authorities. *Briefing papers. House of Commons Library, London*.

Sandford M (2019b) Money talks: The finances of English Combined Authorities. *Local Economy* 34(2): 106-122.

Sandford M (2022) Devolution to local government in England. *Briefing paper.*

Saneinejad S and Roorda MJ (2009) Application of sequence alignment methods in clustering and analysis of routine weekly activity schedules. *Transportation Letters-the International Journal of Transportation Research* 1(3): 197-211.

Schneider F, Ton D, Zomer LB, et al. (2021) Trip chain complexity: a comparison among latent classes of daily mobility patterns. *Transportation* 48(2): 953-975.

Sdoukopoulos A, Pitsiava-Latinopoulou M, Basbas S, et al. (2019) Measuring progress towards transport sustainability through indicators: Analysis and metrics of the main indicator initiatives. *Transportation Research Part D-Transport and Environment* 67: 316-333.

Sengupta U, Rauws WS and de Roo G (2016) Planning and complexity: Engaging with temporal dynamics, uncertainty and complex adaptive systems. *Environment and Planning B-Planning & Design* 43(6): 970-974.

Shaheen S, Cohen A, Dowd MK, et al. (2019) A Framework for Integrating Transportation into Smart Cities.

Shaw R, Kim Y-k and Hua J (2020) Governance, technology and citizen behavior in pandemic: Lessons from COVID-19 in East Asia. *Progress in Disaster Science* 6: 100090.

Shou ZY and Di X (2018) Similarity analysis of frequent sequential activity pattern mining. *Transportation Research Part C-Emerging Technologies* 96: 122-143.

Silva AE, Liu L, Kwon HR, et al. (2021) What's new in urban data analytics? In: Rae A and Wong C (eds) *Applied Data Analysis for Urban Planning and Management*.

Singleton A and Arribas-Bel D (2021) Geographic data science. *Geographical Analysis* 53(1): 61-75.

Sjoman M, Ringenson T and Kramers A (2020) Exploring everyday mobility in a living lab based on economic interventions. *European Transport Research Review* 12(1).

Skrimizea E, Haniotou H and Parra C (2019) On the 'complexity turn' in planning: An adaptive rationale to navigate spaces and times of uncertainty. *Planning Theory* 18(1): 122-142.

Smart London (2018) *Introducing the Smarter London Together Report Card — charting smart city steps openly*. Available at: https://smartlondon.medium.com/introducing-the-smarter-london-together-report-cards-charting-smart-city-steps-openly-eefc7643eda0.

Smart London (2020) *City-wide data in London: pandemic response & recovery (Part 1)*. Available at: https://smartlondon.medium.com/city-wide-data-in-london-pandemic-response-recovery-part-1-13c25efbac43.

Smith A, Stirling A and Berkhout F (2005) The governance of sustainable socio-technical transitions. *Research policy* 34(10): 1491-1510.

Song Y, Ren SY, Wolfson J, et al. (2021) Visualizing, clustering, and characterizing activity-trip sequences via weighted sequence alignment and functional data analysis. *Transportation Research Part C-Emerging Technologies* 126.

South Yorkshire Mayoral Combined Authority (2021) *Our Projects*. Available at: https://southyorkshire-ca.gov.uk/explore/our-projects (Accessed 12/2022)

Sovacool BK, Del Rio DF and Griffiths S (2020) Contextualizing the Covid-19 pandemic for a carbon-constrained world: Insights for sustainability transitions, energy justice, and research methodology. *Energy Research & Social Science* 68.

Spielhofer T, Hahne AS, Reuter C, et al. (2019) Social Media Use in Emergencies of Citizens in the United Kingdom. *ISCRAM.*

Stilgoe J (2018) Machine learning, social learning and the governance of self-driving cars. *Social Studies of Science* 48(1): 25-56.

Studer M (2013) WeightedCluster library manual: A practical guide to creating typologies of trajectories in the social sciences with R.

Studer M (2018) Divisive property-based and fuzzy clustering for sequence analysis. *Sequence analysis and related approaches*. Springer, Cham, pp.223-239.

Studer M and Ritschard G (2016) What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society Series a-Statistics in Society* 179(2): 481-511.

Su RX, McBride EC and Goulias KG (2020) Pattern recognition of daily activity patterns using human mobility motifs and sequence analysis. *Transportation Research Part C-Emerging Technologies* 120.

Sudmant A, Viguie V, Lepetit Q, et al. (2021) Fair weather forecasting? The shortcomings of big data for sustainable development, a case study from Hubballi-Dharwad, India. *Sustainable Development* 29(6): 1237-1248.

Sung J and Monschauer Y (2020) Changes in transport behaviour during the Covid-19 crisis. *IEA.* 2020.

Šurdonja S, Giuffrè T and Deluka-Tibljaš A (2020) Smart mobility solutions– necessary precondition for a well-functioning smart city. *Transportation research procedia* 45: 604-611.

Szczepanek R (2020) Analysis of pedestrian activity before and during COVID-19 lockdown, using webcam time-lapse from Cracow and machine learning. *Peerj* 8.

Taecharungroj V and Mathayomchan B (2020) The big picture of cities: Analysing Flickr photos of 222 cities worldwide. *Cities* 102.

Team RC (2013) R: A language and environment for statistical computing. Vienna, Austria.

Tees Valley (2021) *Tees Valley's Strategic Economic Plan*. https://teesvalley-ca.gov.uk/sep/ (Accessed 12/2022)

Thakuriah P, Tilahun NY and Zellner M (2017) Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery.

Seeing Cities through Big Data: Research, Methods and Applications in Urban Informatics. DOI: 10.1007/978-3-319-40902-3_2. 11-45.

Toh CK, Sanguesa JA, Cano JC, et al. (2020) Advances in smart roads for future smart cities. *Proceedings of the Royal Society A* 476(2233): 20190439.

Tomaszewska EJ and Florea A (2018) Urban smart mobility in the scientific literature—bibliometric analysis. *Engineering Management in Production and Services* 10(2): 41-56.

Townsend A (2019) Combined Authorities for more sub-regions?–Learning the adverse lessons from England beyond the metropolitan conurbations. *Local Economy* 34(2): 123-138.

Transport for London (2011) *Travel in London, Supplementary Report: London Travel Demand Survey (LTDS)*. Available at: https://www.clocs.org.uk/wp-content/uploads/2014/05/london-travel-demand-survey-2011.pdf.

Transport for London (2017) *Transport Classification of Londoners*. Available at: https://content.tfl.gov.uk/transport-classification-of-londoners-presenting-the-segments.pdf.

Transport for London (2020a) *Financial impact of COVID-19 and Government support package*. Available at: https://tfl.gov.uk/info-for/investors/announcements.

Transport for London (2020b) London Travel Demand Survey.

Transport for London (2020c) *TfL announces plan to help London travel safely and sustainably*. Available at: https://tfl.gov.uk/info-for/media/press-releases/2020/may/tfl-announces-plan-to-help-london-travel-safely-and-sustainably.

Transport for London (2020d) *TfL announces plans to make walking and cycling at Lambeth Bridge safer and easier*. Available at: https://tfl.gov.uk/info-for/media/press-releases/2020/march/tfl-announces-plans-to-make-walking-and-cycling-at-lambeth-bridge-safer-and-easier.

Transport for London (2020e) *Travel in London (Report 13)*. Available at: https://content.tfl.gov.uk/travel-in-london-report-13.pdf.

Transport for London (2021) *Travel in London (Report 14)*. Available at: https://content.tfl.gov.uk/travel-in-london-report-14.pdf.

Transport for London (2022) *TfL Statement - TfL Funding Update*. Available at: https://tfl.gov.uk/info-for/media/press-releases/2022/february/tfl-statement---tfl-funding-update-18th-februa (accessed 02/2022).

Uberoi E (2021) *Combined authority mayoral elections in May 2021.* Available at: https://researchbriefings.files.parliament.uk/documents/CBP-9237/CBP-9237.pdf (accessed 12/2022).

Van Eck NJ and Waltman L (2009) VOSviewer: A Computer Program for Bibliometric Mapping. *Proceedings of Issi 2009 - 12th International Conference of the International Society for Scientometrics and Informetrics, Vol 2* 2: 886-897.

Vrscaj D, Nyholm S and Verbong GPJ (2021) Smart mobility innovation policy as boundary work: identifying the challenges of user involvement. *Transport Reviews* 41(2): 210-229.

Walker WE, Marchau VA and Kwakkel JH (2019) Dynamic Adaptive Planning (DAP). *Decision Making under Deep Uncertainty*. Springer, Cham, pp.53-69.

Ward JH (1963) Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* 58(301): 236-&.

Wallenti G (2020) *UNIGIS module: Spatial Simulation.* Available at: https://unigis-salzburg.github.io/Opt_Spatial-Simulation/index.html (accessed 12/2022)

Wells P, Abouarghoub W, Pettit S, et al. (2020) A socio-technical transitions perspective for assessing future sustainability following the COVID-19 pandemic. *Sustainability: Science, Practice and Policy* 16(1): 29-36.

West Midlands Combined Authority (2022) *Plans to give greater powers to the West Midlands welcomed as Levelling Up strategy is published*. Available at: https://www.wmca.org.uk/news/plans-to-give-greater-powers-to-the-west-midlands-welcomed-as-levelling-up-strategy-is-published/.

West Yorkshire Combined Authority (2021) *West Yorkshire Combined Authority.* Available at: https://www.westyorks-ca.gov.uk (Accessed 12/2022)

White PR (2016) *Public transport: its planning, management and operation.* Taylor & Francis.

Whitelaw S, Mamas MA, Topol E, et al. (2020) Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digital Health* 2(8): E435-E440.

Wilson C (1998a) Analysis of travel behavior using sequence alignment methods. *Transportation Research Record* 1645(1): 52-59.

Wilson WC (1998b) Activity pattern analysis by means of sequence-alignment methods. *Environment and Planning A* 30(6): 1017-1038.

Winkler R, Klawonn F, Höppner F, et al. (2010) Fuzzy cluster analysis of larger data sets. *Scalable Fuzzy Algorithms for Data Management and Analysis: Methods and Design*. IGI Global, pp.302-331.

Wissel BD, Van Camp PJ, Kouril M, et al. (2020) An interactive online dashboard for tracking COVID-19 in US counties, cities, and states in real time. *Journal of the American Medical Informatics Association* 27(7): 1121-1125.

Wlezien C (2017) Public Opinion and Policy Representation: On Conceptualization, Measurement, and Interpretation. *Policy Studies Journal* 45(4): 561-582.

Wolfswinkel JF, Furtmueller E and Wilderom CPM (2013) Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems* 22(1): 45-55.

Woods E, Rodriguez Labastida R, Citron R, et al. (2017) UK Smart Cities Index 2017. *Commissioned by Huawei from Navigant Consulting, Inc., http://e.*

huawei. com/uk/special_topic/solution/smart_cities_index_2017, Downloaded 12(3): 18.

Xianyu JC, Rasouli S and Timmermans H (2017) Analysis of variability in multi-day GPS imputed activity-travel diaries using multi-dimensional sequence alignment and panel effects regression models. *Transportation* 44(3): 533-553.

Xu G, Xiu TY, Li X, et al. (2021) Lockdown induced night-time light dynamics during the COVID-19 epidemic in global megacities. *International Journal of Applied Earth Observation and Geoinformation* 102.

Xu L and Kwan MP (2020) Mining sequential activity-travel patterns for individual-level human activity prediction using Bayesian networks. *Transactions in Gis* 24(5): 1341-1358.

Xu L and McArdle G (2018) Internet of too many things in smart transport: The problem, the side effects and the solution. *Ieee Access* 6: 62840-62848.

Xu R and Wunsch DC (2010) Clustering algorithms in biomedical research: a review. *IEEE reviews in biomedical engineering* 3: 120-154.

Yang KF (2020) Unprecedented Challenges, Familiar Paradoxes:COVID-19 and Governance in a New Normal State of Risks. *Public Administration Review* 80(4): 657-664.

Yatskiv I, Nathanail E, Savrasovs M, et al. (2018) Assessing Knowledge Level of Stakeholders on Transport Interchange Design and Operation. *Transport* 33(3): 793-800.

Yousif W and Fox M (2018) *A Transportation Ontology for Global City Indicators (ISO 37120)*.

Zhang N, Jia W, Wang PH, et al. (2021) Changes in local travel behaviour before and during the COVID-19 pandemic in Hong Kong. *Cities* 112.

Zhang YH, Zhang AM and Wang JE (2020) Exploring the roles of high-speed train, air and coach services in the spread of COVID-19 in China. *Transport Policy* 94: 34-42.

Zhang YP, Liu L and Wang H (2019) A new perspective on the temporal pattern of human activities in cities: The case of Shanghai. *Cities* 87: 196-204.

Appendix

Appendix A: Characteristics of the metropolitan areas

Table A-1: Population, area, density and GVA of the metropolitan areas

Metropolis	Areas (km2)	Population[1]	Density	Total GVA[2]
Greater London	1569	8908081	5678	431164
West Midlands	902	2916458	3235	66667
Greater Manchester	1276	2812569	2204	66413
West Yorkshire	2029	2320214	1143	50766[3]
Liverpool City Region	726	1551497	2138	32030
Sheffield City Region	1552	1402918	904	25991
North of Tyne	5222	1157170	222	18863
West of England	958	938155	980	29295
Cambridgeshire and Peterborough	3396	852523	251	24463
North East	2576	816000	317	37871[3]
Tees Valley	795	674284	848	13122

Note: Data source: ONS

1. Population estimates are sourced from population estimates for UK, 2018

2. GVA(B) in current prices - a balanced measure of regional GVA, 2017

3. GVA(B) in current prices - a balanced measure of regional GVA, 2016

Appendix B: Preliminary results of sequence analysis

Appendix B-1: Dissimilarity measure selection

The four distance matrices we compared are: 1) unit cost optimal matching (OM) as the reference metric (OM), 2) transition-based OM (OMtrate), 3) Localised OM that emphasises adjacent states (OMloc), and 4) Dynamic Hamming distance that considers time-dependent transition rate between two states (DHD). Four distance measures were applied to the simplest sequence type that contains the 9-state activity-trip (i.e., eight activities and one trip type) in the preliminary analysis. We selected 10% of activity-trip weekday sequences from 2015 to 2019 randomly. In total, 4814 sequences were selected and weighted (using the interim expansion factor for the weekday sample). We compared the distances by computation time and clustering quality. We applied the most used Ward clustering method to group the selected sequences.

For unit cost OM distance, the elapsed time is 9.5 minutes. It takes 9.7 minutes to calculate transition-based OM dissimilarities. Regarding Localised OM distance, the computational time is 2.1 hours, which is much longer than the other metrics. The elapsed time of the DHD method is 10.2 seconds, which is the fastest among all distances. Agglomerative coefficients (AC) that describe the clustering structure of Ward clustering results are all around 0.999. When AC is higher, a better cluster result and dendrogram are more likely to be obtained (Bhandari and Pahwa, 2020). AC of clustering result based on transition-based OM is the highest while that of DHD is the lowest. The DHD-based clustering result is substantially different from the three OM-based dendrograms. This is probably because the DHD metric does not use INDEL costs. The INDEL refers to the insertion or deletion of an element that can cause a one-position change in OM (Kim, 2018).

The total inertia of the dendrogram, as seen in the “height” object, was used to decide the optimal number of clusters. Previous studies such as Hafezi et al. (2017), Jiang et al. (2012) and Allahviranloo et al. (2017) found eight clusters of activity-travel patterns. We chose the optimal number around eight, despite the three clusters had the best validity indices. According to the inertia plots in figure B-1, we decided to use seven clusters for comparison.

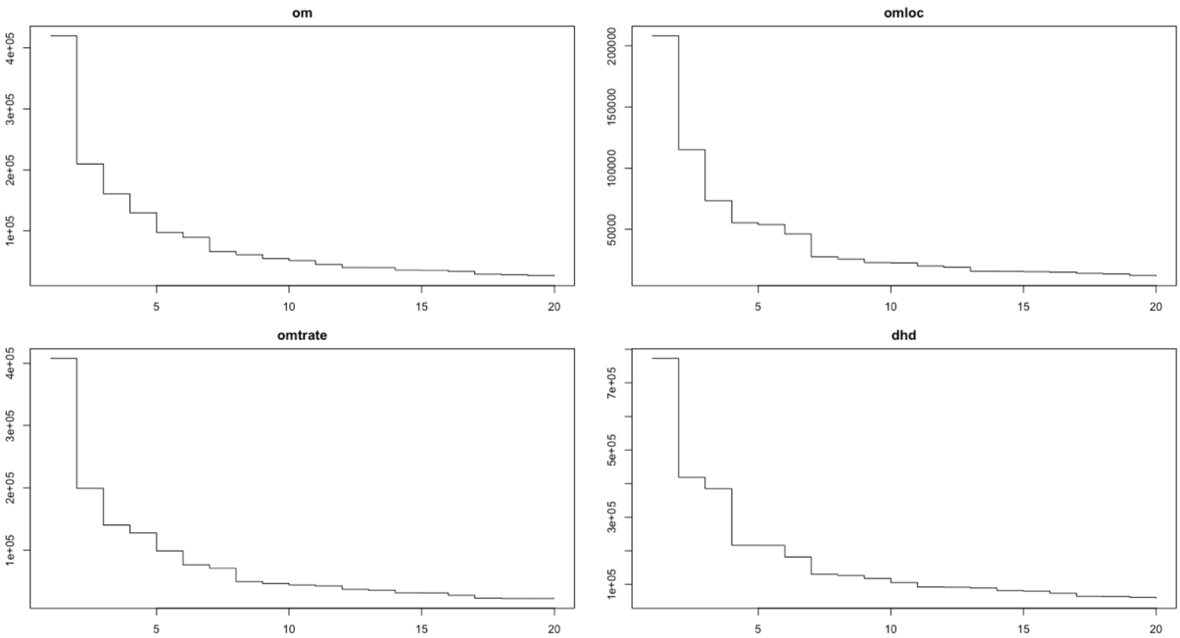

Figure B-1: Plots of total inertia

We first compared the three widely used internal validation indices – the average silhouette widths, the Dunn index and connectivity - of four clustering results. The result showed that the DHD scored best in the Dunn index and OMtrate had the best connectedness and scores in average silhouette widths.

We plotted distributions of different activities and the ten most frequent sequences (see Figure B-2). We also compared the mean time spent in each state, and entropy of state distribution by the time of clustering results based on the four metrics. Cluster 3 in the DHD-based result is similar to Cluster 7 in the rest clustering results. Likewise, Cluster 7 in the DHD-based result corresponds to Cluster 3 in the other clustering results. Compared to clustering results of OM, those of OMloc and OMtrate identify a different pattern in Cluster 6, which contains shorter hours working activities. Cluster 6 of OM-based and DHD-based results are recreational activities in midday/afternoon.

Although the DHD has the shortest computation time and the best Dunn index score, it does not take INDEL into account when calculating distance. Thus, the difference in activity type is less captured. Considering both computation time and clustering results, we decided to use the OMtrate distance for further analysis in Chapter 4.

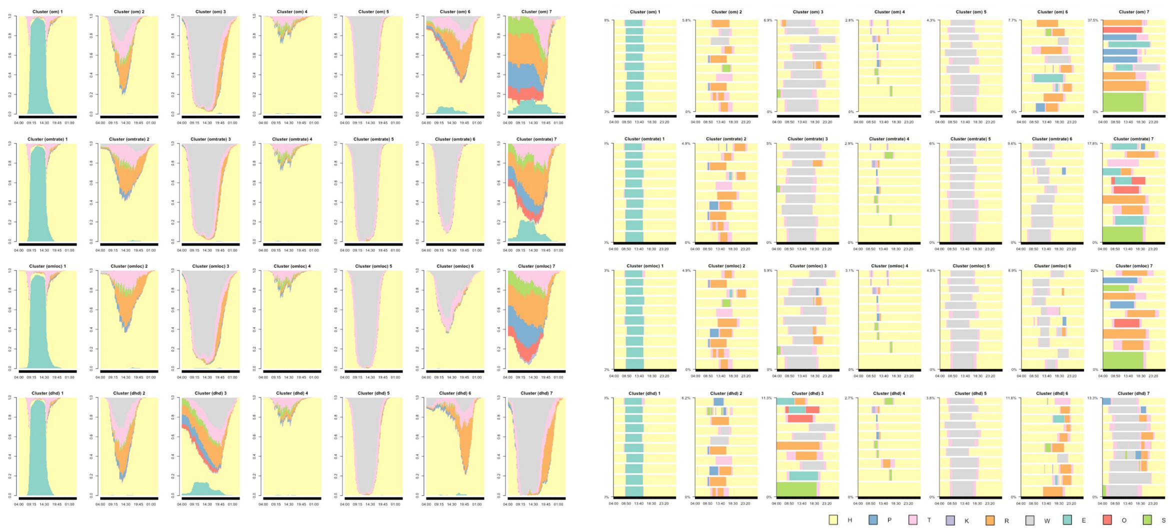

Figure B-2: Distributions of different activities (left) & 10 most frequent sequences (right)

Appendix B-2: Clustering method selection

The four clustering methods we compared are 1) hierarchical clustering, 2) FANNY, 3) NEFRC, and 4) NEFRC with noise clustering. Four clustering methods were applied to the 5% simplest sequence type that contains 9-state activity-trip. 2405 sequences from 2015 to 2019 were selected for comparison. As shown in Table B-1, we compared the clustering methods by computation times and internal clustering qualities (i.e., Silhouette index). FANNY clustering method had the best quality and fastest computation time among fuzzy clustering methods. Both Ward and FANNY methods produced relatively good results in a shorter time. However, the FANNY method allows more flexibility in the clustering, which can provide a more adaptive understanding. Thus, we use the FANNY method for further clustering in Chapter 4.

Table B-1: Clustering methods comparison

Clustering method	Quality (3 clusters)	Quality (4 clusters)	Computation time
Ward	0.5456	0.5038	fast
FANNY	0.5224	0.5291	medium
NEFRC	0.4874	0.4576	slow
NEFRC.noise	0.4805	0.426	slow

Appendix B-3: Yearly and weekday/weekend differences

A comparison of main complexity characteristics and representative patterns from 2015 to 2019 indicates some similarities in trip-activities of Londoners in the previous years. We measured the complexity of trip activity trajectories through four indices – transition, longitudinal entropy, turbulence, and complexity. We presented the mean of these indices for each year in Table B-2.

As shown in the transition index, the number of activity changes on weekdays ranges from 5.7 to 6.0. State changes in weekends' sequences are slightly less, ranging from 5.5 to 5.9. The entropy level of the individual trip-activity sequences ranges from 0.31 to 0.32 on weekdays and from 0.26 to 0.27 at weekends. The turbulence index is stable in each year. Most years witnessed a within-sequence turbulence of 0.026 on weekdays and 0.025 at weekends. Showing overall complexity level, the means of complexity index are mainly 0.076 on weekdays and 0.068 at the weekends.

The complexity of the individuals' daily sequences was similar every year but varied on weekdays and weekends. Thus, we analysed all sequences from 2015 to 2019 to find main patterns and spilt these sequences into weekdays' and weekends' for pattern analysis in Chapter 4.

Table B-2: Complexity indices of trip-activity sequences

Days	Year	Number	Transitions	Entropy	Turbulence	Complexity
Weekdays	2015	9934	6.021	0.321	0.027	0.079
	2016	9630	5.847	0.320	0.026	0.077
	2017	9554	5.683	0.317	0.026	0.076
	2018	9423	5.678	0.315	0.026	0.076
	2019	9601	5.853	0.314	0.026	0.076
Weekends	2015	3856	5.859	0.272	0.026	0.072
	2016	3592	5.681	0.271	0.025	0.071
	2017	3426	5.604	0.267	0.025	0.069
	2018	3337	5.531	0.261	0.024	0.068
	2019	3851	5.558	0.261	0.025	0.068

The number of weekdays' clusters varied in the range of seven to nine in recent years (2015-2019), as shown in Table B-2. A comparison of the representative sequences in different years can show the similarities in the past years. In the previous years, most Londoners spent their time at home, work, and education. Activities usually took place in the daytime.

Day work activities can be found in all the years. Based on the working and home locations, the three main spatiotemporal patterns are: 1) Inner London day work (and home), 2) Outer London day work (and home), and 3) Inner London day work (and Outer London home).

Day education activities are identified in most of the years. The day education patterns can be divided into two groups: 1) Inner London day education, and 2) Outer London day education.

Table B-3: Representative sequences of yearly clusters on weekdays

Year	Cluster	Trip-activity Reference	Location reference	Pattern
2015	8	H/95-TV/4-R/50-TV/6-H/133	OR/288	Outer London midday recreation
		H/288	OR/288	Outer London stay-at-home
		H/47-TB/8-W/112-TB/10-H/111	OR/47-IW/120-OR/121	Inner London day work (living in Outer London)
		H/51-TW/1-E/85-TW/1-H/150	OR/51-OC/86-OR/151	Outer London day education
		H/47-TV/3-W/105-TV/3-H/130	OR/47-OW/108-OR/133	Outer London day work
		H/51-TB/6-W/108-TB/6-H/117	IR/51-IW/114-IR/123	Inner London day work
		H/288	IR/288	Inner London stay-at-home
		H/288	ER/288	External London stay-at-home
2016	7	H/47-TB/9-W/111-TB/9-H/112	OR/47-IW/120-OR/121	Inner London day work (living in Outer London)
		H/288	OR/288	Outer London stay-at-home
		H/53-TW/1-E/83-TW/1-H/150	OR/53-OC/84-OR/151	Outer London day education
		H/44-TV/1-K/1-TV/1-W/108-TV/2-H/131	OR/46-OW/109-OR/133	Outer London day work
		H/52-TB/6-W/103-TB/9-H/118	IR/52-IW/109-IR/127	Inner London day work
		H/288	IR/288	Inner London stay-at-home
		H/53-TW/1-E/83-TW/1-H/150	IR/53-IC/84-IR/151	Inner London day education
2017	7	H/288	OR/288	Outer London stay-at-home
		H/47-TV/2-W/103-TV/3-H/133	OR/47-OW/105-OR/136	Outer London day work
		H/50-TW/1-E/86-TW/1-H/150	OR/50-OC/87-OR/151	Outer London day education
		H/47-TB/9-W/111-TB/9-H/112	OR/47-IW/120-OR/121	Inner London day work (living in Outer London)
		H/288	IR/288	Inner London stay-at-home

		H/53-TB/6-W/105-TB/6-H/118	IR/53-IW/111-IR/124	Inner London day work
		H/53-TW/1-E/84-TW/1-H/149	IR/53-IC/85-IR/150	Inner London day education
2018	8	H/288	OR/288	Outer London stay-at-home
		H/47-TB/6-W/108-TB/6-H/121	OR/47-IW/114-OR/127	Inner London day work (living in Outer London)
		H/47-TB/4-W/104-TB/4-H/129	OR/47-OW/108-OR/133	Outer London day work
		H/53-TV/1-E/83-H/151	OR/53-OC/84-OR/151	Outer London day education
		H/185-W/1-H/102	IR/185-IH/1-IR/102	Inner London stay-at-home
		H/47-TB/6-W/108-TB/6-H/121	IR/47-IW/114-IR/127	Inner London day work
		H/53-TW/1-E/83-TW/1-H/150	IR/53-IC/84-IR/151	Inner London day education
		H/95-TW/1-S/1-H/191	ER/95-ES/2-ER/191	External London mixed activities
2019	9	H/288	OR/288	Outer London stay-at-home
		H/101-TB/3-R/27-TB/3-H/154	OR/101-OP/30-OR/157	Outer London day recreation
		H/47-TB/6-W/102-TB/6-H/127	OR/47-OW/108-OR/133	Outer London day work
		H/44-TB/12-W/108-TB/12-H/112	OR/44-IW/120-OR/124	Inner London day work (living in Outer London)
		H/51-TW/1-E/85-TW/1-H/150	OR/51-OC/86-OR/151	Outer London day education
		H/95-TW/1-H/192	IR/95-IG/1-IR/192	Inner London stay-at-home
		H/136-TV/1-K/1-TV/1-H/149	ER/136-EC/2-ER/150	External London mixed activities
		H/49-TB/6-W/109-TB/6-H/118	IR/49-IW/115-IR/124	Inner London day work
		H/53-TW/1-E/83-TW/1-H/150	IR/53-IC/84-IR/151	Inner London day education

The cluster numbers of weekends' sequences are mostly five, as shown in Table B-3. The main activities at the weekends are home-based activities, recreation, and work. Most activities occurred in the daytime while activities such as recreation, personal business and working can last overnight.

A large proportion of Londoners spent their weekends at home. Based on the home locations, we identified Outer London stay-at-home, Inner London stay-at-home, and External London stay-at-home groups. Additionally, day recreation or day work patterns can be found in most of the years. The recreational activities can occur at midday, afternoon, and overnight while the working activities mainly took place in the daytime.

Table B-4: Representative sequences of yearly clusters on weekends

Year	Clusters	Trip-activity Reference	Location reference	Activity Pattern
2015	5	H/288	OR/288	Outer London stay-at-home
		H/74-TV/4-R/75-TV/5-H/130	OR/288	Outer London day recreation/work
		H/107-TV/1-K/1-TV/1-H/178	ER/107-EO/2-ER/179	External London stay-at-home
		H/95-TV/3-R/11-TW/1-R/41-TB/4-H/133	IR/95-IP/56-IR/137	Inner London day recreation/work
		H/288	IR/288	Inner London stay-at-home
2016	5	H/131-TW/1-H/156	OR/288	Outer London stay-at-home
		H/107-TV/2-R/56-TV/2-H/121	OR/288	Outer London day recreation
		H/47-TV/6-W/102-TV/6-H/127	OR/47-ER/108-OR/133	Outer London day work
		H/95-TB/6-R/42-TB/6-H/139	IR/95-IP/48-IR/145	Inner London day recreation/work
		H/107-TW/1-H/180	IR/107-IS/1-IR/180	Inner London stay-at-home
2017	5	H/104-TV/3-R/63-TV/3-H/115	OR/288	Outer London day recreation
		H/288	OR/288	Outer London stay-at-home
		H/41-TV/5-W/91-TV/6-H/145	OR/41-OP/96-OR/151	Overnight activities
		H/83-TW/2-H/203	IR/83-IS/1-IR/204	Inner London stay-at-home
		H/95-TW/9-R/51-TW/9-H/124	IR/95-IP/60-IR/133	Inner London day work/recreation
2018	5	H/107-TW/2-R/46-TW/2-H/131	OR/288	Outer London day recreation
		H/288	OR/288	Outer London stay-at-home
		H/47-TB/3-W/93-TB/3-H/142	OR/47-OP/96-OR/145	Outer London day work
		H/288	IR/288	Inner London stay-at-home
		H/77-TW/1-S/2-TW/1-H/207	ER/77-ES/3-ER/208	External London stay-at-home

2019	5	H/288	OR/288	Outer London stay-at-home
		H/89-TW/1-R/77-TW/1-H/120	OR/288	Outer London day recreation
		H/95-TB/4-R/48-TB/4-H/137	IR/95-IP/52-IR/141	Inner London day recreation
		H/83-TW/1-H/204	IR/83-IS/1-IR/204	Inner London stay-at-home
		H/113-TW/1-S/2-TW/1-H/171	ER/113-ES/3-ER/172	External London stay-at-home

Appendix C: Detailed tables for weekdays' clusters

Table C-1: Dirichlet regression of cluster membership (2015-2019 weekdays)

Cluster (baseline: Group #8)		Group #1		Group #2		Group #3		Group #4		Group #5		Group #6		Group #7		Group #9		Group #10		Group #11	
Attributes	Categories	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*
Personal characteristics																					
Gender	M	0.016	0.388	-0.003	0.867	-0.004	0.849	0.01	0.582	0.007	0.692	0.014	0.435	0.009	0.627	0.02	0.272	-0.007	0.696	0.007	0.701
	F	ref																			
Age	A1	ref																			
	A2	-0.854	<0.001***	-0.365	<0.001***	-0.106	<0.05*	-0.142	<0.01**	-0.151	<0.01**	-0.196	<0.001***	-0.163	<0.001***	-0.046	0.315	-0.009	0.848	-0.182	<0.001***
	A3	-0.903	<0.001**	-0.438	<0.001**	-0.153	<0.01**	-0.192	<0.001**	-0.217	<0.001**	-0.238	<0.001**	-0.213	<0.001**	-0.052	0.358	-0.035	0.534	-0.224	<0.001**

			*		*				*		*		*		*						*
	A4	-0.936	<0.001***	-0.463	<0.001***	-0.155	<0.001**	-0.279	<0.001***	-0.23	<0.001***	-0.367	<0.001***	-0.281	<0.001***	-0.09	0.129	-0.025	0.669	-0.258	<0.001***
	A5	-1.008	<0.001***	-0.573	<0.001***	-0.235	<0.001***	-0.412	<0.001***	-0.32	<0.001***	-0.524	<0.001***	-0.377	<0.001***	-0.172	<0.05*	-0.04	0.556	-0.372	<0.001***
Ethic group	E1	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref	ref
	E2	-0.041	0.252	-0.02	0.587	-0.015	0.676	-0.052	0.153	-0.012	0.735	-0.058	0.112	-0.009	0.802	-0.041	0.265	-0.009	0.808	-0.048	0.193
	E3	-0.025	0.312	-0.068	<0.01**	-0.02	0.422	-0.108	<0.001***	0	0.995	-0.107	<0.001***	-0.011	0.644	-0.107	<0.001***	-0.015	0.54	-0.071	<0.01**
	E4	0.021	0.47	0.206	<0.001***	0.085	<0.01**	0.16	<0.001***	0.101	<0.001***	0.132	<0.001***	0.095	<0.01**	0.213	<0.001***	0.037	0.203	0.123	<0.001***
Dr	D	ref																			

ivinglicenseholder	ND	0.116	<0.001***	-0.031	0.2	0.028	0.243	0.004	0.879	0.019	0.442	-0.014	0.573	0.005	0.833	0.045	<0.1.	0.008	0.743	0.02	0.413
Oystercardholder	B	ref																			
	NB	0.024	0.27	0.082	<0.001***	0.001	0.979	0.067	<0.01**	0.005	0.83	0.074	<0.001***	-0.002	0.941	0.078	<0.001***	0	0.985	0.046	<0.05*
Occupation	O1	ref																			
	O2	-0.236	<0.001***	-0.264	<0.001***	-0.396	<0.001***	-0.383	<0.001***	-0.484	<0.001***	-0.444	<0.001***	-0.5	<0.001***	-0.146	<0.001***	-0.073	<0.05*	-0.273	<0.001***
	O3	0.248	<0.001***	0.134	<0.01**	-0.419	<0.001***	-0.348	<0.001***	-0.509	<0.001***	-0.426	<0.001***	-0.515	<0.001***	-0.029	0.556	-0.052	0.288	-0.167	<0.001***
	O4	-0.613	<0.001***	-0.694	<0.001***	-0.944	<0.001***	-0.922	<0.001***	-1.023	<0.001***	-0.987	<0.001***	-1.035	<0.001***	-0.355	<0.001***	-0.286	<0.001***	-0.696	<0.001***

	O5	-0.547	<0.001***	-0.667	<0.001***	-0.895	<0.001***	-0.874	<0.001***	-0.97	<0.001***	-0.925	<0.001***	-0.964	<0.001***	-0.295	<0.001***	-0.248	<0.01**	-0.664	<0.001***
	O6	-0.713	<0.001***	-0.822	<0.001***	-1.032	<0.001***	-1.015	<0.001***	-1.12	<0.001***	-1.027	<0.001***	-1.089	<0.001***	-0.657	<0.001***	-0.332	<0.001***	-0.8	<0.001***
	O7	-0.785	<0.001***	-0.769	<0.001***	-1.114	<0.001***	-1.003	<0.001***	-1.182	<0.001***	-1.052	<0.001***	-1.194	<0.001***	-0.47	<0.001***	-0.412	<0.001***	-0.813	<0.001***
	O8	-0.331	<0.01**	-0.369	<0.001***	-0.596	<0.001***	-0.554	<0.001***	-0.674	<0.001***	-0.612	<0.001***	-0.669	<0.001***	-0.144	0.152	-0.111	0.267	-0.355	<0.001***
Health condition	H	ref																			
	HP	-0.058	0.144	-0.042	0.291	-0.037	0.35	-0.061	0.127	-0.038	0.341	-0.067	<0.1.	-0.037	0.35	-0.05	0.209	-0.023	0.56	-0.039	0.332
	HM	0.08	0.447	0.14	0.188	0.09	0.395	0.09	0.395	0.10	0.31	0.06	0.531	0.09	0.358	0.27	<0.01	0.01	0.91	0.10	0.338

		1		2		2		2		9		8		9		4	**	2		3	
Household Characteristics																					
Accessible vehicles	C0	ref																			
	C1	-0.036	0.123	-0.366	<0.001***	-0.081	<0.001***	-0.358	<0.001***	-0.104	<0.001***	-0.375	<0.001***	-0.129	<0.001***	-0.389	<0.001***	-0.04	<0.1.	-0.207	<0.001***
	C2	0.031	0.276	-0.603	<0.001***	-0.061	<0.05*	-0.565	<0.001***	-0.111	<0.001***	-0.594	<0.001***	-0.176	<0.001***	-0.615	<0.001***	-0.024	0.404	-0.305	<0.001***
Household income	IL	-0.108	<0.001***	-0.073	<0.05*	-0.029	0.319	-0.15	<0.001***	-0.054	<0.1.	-0.203	<0.001***	-0.09	<0.01**	-0.099	<0.001***	-0.02	0.496	-0.073	<0.05*
	IM	-0.042	<0.05*	-0.078	<0.001***	-0.039	<0.1.	-0.108	<0.001***	-0.061	<0.01**	-0.148	<0.001***	-0.094	<0.001***	-0.098	<0.001***	-0.015	0.485	-0.059	<0.01**
	IH	ref																			
Household	H1	ref																			
	H2	-0.	0.49	0.0	<0.1	0.0	0.32	0.0	<0.0	0.0	<0.1	0.1	<0.0	0.0	<0.1	0.0	0.59	0.0	0.75	0.0	<0.1

structure		017	2	48	.	25	7	91	01***	49	.	2	01***	45	.	13	9	08	6	45	.
	H3	0.037	0.282	0.044	0.2	0.018	0.606	0.116	<0.001***	0.021	0.537	0.128	<0.001***	0.025	0.461	0.084	<0.05*	0.021	0.544	0.068	<0.05*
	H4	0.091	<0.05*	0.091	<0.05*	0.048	0.273	0.167	<0.001***	0.076	<0.1.	0.183	<0.001***	0.066	0.133	0.174	<0.001***	0.024	0.593	0.109	<0.05*
	H5	0.059	<0.1.	0.022	0.501	0.06	<0.1.	0.071	<0.05*	0.071	<0.05*	0.089	<0.01**	0.048	0.141	0.02	0.55	0.023	0.491	0.045	0.169
	H6	0.07	0.149	0.194	<0.001***	0.085	<0.1.	0.231	<0.001***	0.103	<0.05*	0.241	<0.001***	0.109	<0.05*	0.207	<0.001***	0.042	0.395	0.16	<0.01**

Note: Number of Observations: 24070, Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table C-2: Proportion of socio-demographic covariates in 11 clusters (2015-2019 weekdays)

Attributes	Categories	All	Group #1	Group #2	Group #3	Group #4	Group #5	Group #6	Group #7	Group #8	Group #9	Group #10	Group #11	Chi-squared P value
Personal Characteristics														
Gender	M	49	51	49	52	52	56	54	57	44	43	46	46	<0.001
	F	51	49	51	48	48	44	46	43	56	57	54	54	
Age	A1	15	81	73	1	1	0	0	0	6	6	7	17	<0.001
	A2	11	16	19	10	11	9	9	8	7	11	11	9	
	A3	36	2	6	48	58	50	68	62	29	34	28	36	
	A4	20	0	1	33	23	33	20	25	20	20	21	23	
	A5	18	0	0	9	7	8	3	5	38	29	33	15	
Ethic group	E1	65	53	48	66	71	64	77	66	68	65	69	69	<0.001
	E2	7	12	12	5	6	5	5	6	6	6	6	4	
	E3	17	21	18	19	10	22	9	20	18	13	17	14	
	E4	12	14	22	10	13	9	9	8	9	16	8	13	
Driving license holder	D	57	4	8	75	67	81	72	83	64	51	64	57	<0.001
	ND	43	96	92	25	33	19	28	17	36	49	36	43	
Oyster card holder	B	53	25	26	70	70	71	68	76	45	49	48	49	<0.001
	NB	47	75	74	30	30	29	32	24	55	51	52	51	

Health condition	H	93	97	96	98	98	98	99	99	86	86	88	92	<0.001
	HP	7	2	3	2	2	2	1	1	13	12	11	7	
	HM	1	1	1	0	0	0	0	0	1	2	1	1	
Occupation	O1	47	1	2	83	86	92	95	95	25	28	32	46	<0.001
	O2	9	0	1	12	9	6	5	4	11	13	14	10	
	O3	22	99	97	3	3	0	0	0	10	13	14	21	
	O4	2	0	0	0	0	0	0	0	5	6	4	2	
	O5	2	0	0	0	0	0	0	0	3	5	2	2	
	O6	13	0	0	1	1	0	0	0	32	23	26	12	
	O7	5	0	0	0	0	0	0	0	12	11	5	5	
	O8	1	0	0	0	0	0	0	0	1	2	2	1	
Household Characteristics													
Accessible household vehicles	C0	32	19	44	22	50	18	55	22	25	51	26	29	<0.001
	C1	44	48	47	40	39	45	37	49	49	40	43	42	
	C2	24	32	9	38	11	36	9	29	26	9	31	29	
Household income	IL	19	22	35	10	9	5	4	5	26	31	22	19	<0.001
	IM	46	48	45	49	47	43	36	36	48	42	50	51	
	IH	35	29	21	41	44	52	60	59	26	26	28	30	
Household structur	H1	39	82	73	35	26	35	21	39	34	31	30	39	<0.001
	H2	30	3	5	37	36	40	40	38	36	26	37	31	
	H3	1	5	9	9	16	7	18	8	7	15	9	9	

e		0												
	H4	6	0	2	2	6	3	6	2	9	12	8	6	
	H5	1 1	7	7	14	11	12	11	9	11	12	12	10	
	H6	4	3	4	4	5	4	4	3	3	4	3	5	

Table C-3: Dirichlet regression of cluster membership (2015-2019 weekends)

Cluster (baseline: Group #1)		Group#2		Group#3		Group#4		Group#5		Group#6		Group#7		Group#8	
Attributes	Categories	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*	*Precision*	*Pr(>|z|)*
Personal Characteristics															
Gender	M	0.006	0.787	-0.014	0.516	-0.01	0.638	-0.015	0.48	-0.004	0.846	-0.004	0.831	-0.009	0.673
	F	ref													
Age	A1	ref													
	A2	0.167	<0.01 **	0.153	<0.01 **	0.052	0.332	0.007	0.892	0.175	<0.01 **	0.145	<0.01 **	0.192	<0.001 ***
	A3	0.15	<0.05 *	0.075	0.252	0.01	0.881	-0.01	0.878	0.129	<0.05 *	0.055	0.401	0.117	<0.1 .
	A4	0.036	0.589	-0.019	0.78	-0.051	0.455	-0.045	0.505	-0.004	0.956	-0.016	0.812	-0.007	0.919
	A5	-0.044	0.568	-0.071	0.356	-0.064	0.402	-0.045	0.551	-0.074	0.331	-0.045	0.561	-0.066	0.387
Ethnic group	E1	ref													
	E2	-0.071	0.101	-0.028	0.512	0.009	0.84	0.003	0.951	-0.058	0.179	0.003	0.951	-0.045	0.303
	E3	-0.176	<0.001 ***	-0.098	<0.001 ***	-0.01	0.713	-0.015	0.604	-0.161	<0.001 ***	-0.051	<0.1 .	-0.142	<0.001 ***
	E4	0.284	<0.001 ***	0.168	<0.001 ***	0.078	<0.05 *	0.051	0.136	0.262	<0.001 ***	0.136	<0.001 ***	0.238	<0.001 ***
Drivi	D	ref													

ng license holder	ND	-0.114	<0.001***	-0.041	0.146	-0.006	0.832	-0.002	0.94	-0.111	<0.001***	-0.002	0.931	-0.086	<0.01**
Oyster card holder	B	ref													
	NB	0.117	<0.001***	0.068	<0.01**	-0.006	0.814	-0.007	0.781	0.096	<0.001***	0.012	0.613	0.075	<0.01**
Occupation	O1	ref													
	O2	-0.039	0.296	-0.048	0.206	-0.011	0.776	-0.002	0.951	-0.047	0.21	-0.047	0.215	-0.056	0.141
	O3	0.084	0.133	-0.032	0.568	-0.01	0.859	-0.01	0.86	0.065	0.247	-0.062	0.274	0.027	0.627
	O4	-0.051	0.446	-0.137	<0.05*	-0.058	0.389	-0.057	0.397	-0.08	0.237	-0.187	<0.01**	-0.139	<0.05*
	O5	0.077	0.416	-0.103	0.282	-0.04	0.681	-0.037	0.703	-0.022	0.819	-0.163	<0.1 .	-0.093	0.333
	O6	-0.487	<0.001***	-0.386	<0.001***	-0.192	<0.001***	-0.125	<0.05 *	-0.493	<0.001***	-0.361	<0.001***	-0.49	<0.001***
	O7	-0.08	0.115	-0.148	<0.01**	-0.108	<0.05*	-0.079	0.118	-0.127	<0.05*	-0.215	<0.001***	-0.175	<0.001***
	O8	0.276	<0.05*	0.004	0.973	-0.036	0.746	-0.035	0.748	0.194	<0.1 .	-0.069	0.535	0.116	0.291
Health condition	H	ref													
	HP	0.016	0.726	-0.004	0.935	-0.027	0.577	-0.017	0.729	-0.001	0.985	-0.016	0.737	-0.013	0.779
	HM	-0.3	<0.05	-0.2	<0.05	-0.1	0.333	-0.0	0.48	-0.3	<0.01	-0.1	0.146	-0.3	<0.05

		39	*	9	*	31		95	2	55	**	97		4	*
Household Characteristics															
Accessible vehicles	C0	ref													
	C1	-0.677	<0.001 ***	-0.303	<0.001 ***	-0.073	<0.01 **	-0.042	0.126	-0.626	<0.001 ***	-0.183	<0.001 ***	-0.534	<0.001 ***
	C2	-1.149	<0.001 ***	-0.543	<0.001 ***	-0.122	<0.001 ***	-0.067	<0.05 *	-1.072	<0.001 ***	-0.324	<0.001 ***	-0.93	<0.001 ***
Household income	IL	-0.29	<0.001 ***	-0.168	<0.001 ***	-0.052	0.128	-0.038	0.262	-0.282	<0.001 ***	-0.11	<0.01 **	-0.253	<0.001 ***
	IM	-0.177	<0.001 ***	-0.097	<0.001 ***	-0.03	0.205	-0.025	0.281	-0.168	<0.001 ***	-0.054	<0.05 *	-0.143	<0.001 ***
	IH	ref													
Household structure	H1	ref													
	H2	0.194	<0.001 ***	0.129	<0.001 ***	0.047	<0.1 .	0.009	0.747	0.194	<0.001 ***	0.085	<0.01 **	0.177	<0.001 ***
	H3	0.222	<0.001 ***	0.156	<0.001 ***	0.063	0.115	0.012	0.767	0.231	<0.001 ***	0.113	<0.01 **	0.227	<0.001 ***
	H4	0.276	<0.001 ***	0.168	<0.01 **	0.062	0.234	0.02	0.695	0.257	<0.001 ***	0.114	<0.05 *	0.233	<0.001 ***
	H5	0.239	<0.001 ***	0.137	<0.001 ***	0.065	<0.1 .	0.017	0.667	0.24	<0.001 ***	0.111	<0.01 **	0.22	<0.001 ***
	H6	0.3	<0.001 ***	0.22	<0.001 ***	0.102	<0.1 .	0.053	0.343	0.308	<0.001 ***	0.171	<0.01 **	0.294	<0.001 ***

Note: Number of Observations: 18062, Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table C-4: Proportion of socio-demographic covariates in 8 clusters (2015-2019 weekends)

Attributes	Categories	All	Group#1	Group#2	Group#3	Group#4	Group#5	Group#6	Group#7	Group#8	Chi-squared P value
Personal Characteristics											
Gender	M	49	49	49	46	49	47	49	55	50	<0.001
	F	51	51	51	54	51	53	51	45	50	
Age	A1	14	15	13	13	15	20	13	9	8	<0.001
	A2	11	8	10	11	12	8	12	16	18	
	A3	38	32	39	39	37	33	48	41	48	
	A4	20	22	21	21	20	20	16	24	17	
	A5	18	24	18	17	17	19	11	10	10	
Ethic group	E1	66	67	65	71	63	67	69	65	70	<0.001
	E2	6	6	6	5	6	6	7	6	5	
	E3	17	19	13	15	20	17	13	19	13	
	E4	11	8	16	9	11	10	11	10	12	
Driving license holder	D	60	62	54	65	61	60	58	64	58	<0.001
	ND	40	38	46	35	39	40	42	36	42	
Oyster card holder	B	54	50	52	47	57	51	58	64	62	<0.001
	NB	46	50	48	53	43	49	42	36	38	
Health condition	H	93	91	91	94	95	93	95	95	97	<0.001
	HP	6	8	8	6	5	6	5	4	3	
	HM	1	1	1	0	1	1	0	0	0	
Occupation	O1	48	41	47	53	47	42	53	66	64	<0.001
	O2	9	10	9	10	10	10	9	9	9	
	O3	20	19	19	17	22	25	20	16	18	
	O4	2	3	3	2	3	2	4	1	2	
	O5	2	2	3	1	1	1	2	1	1	
	O6	12	19	12	11	12	14	7	5	5	
	O7	5	6	6	6	5	5	4	2	1	
	O8	1	1	1	0	0	1	1	1	1	

Household Characteristics											
Accessible household vehicles	C0	31	22	47	22	22	21	50	23	49	<0.001
	C1	44	46	42	47	44	46	40	41	40	
	C2	26	33	11	31	34	33	10	35	11	
Household income	IL	17	19	21	13	17	16	17	11	14	<0.001
	IM	47	48	43	46	49	47	42	52	46	
	IH	36	33	36	41	34	37	41	37	40	
Household structure	H1	39	42	35	39	41	50	33	35	27	<0.001
	H2	32	34	29	38	32	30	32	35	32	
	H3	10	7	14	8	9	5	15	9	19	
	H4	6	6	8	5	4	4	6	3	7	
	H5	10	9	11	6	10	7	9	12	11	
	H6	4	3	4	4	4	3	5	5	4	