Technische Universität München
TUM School of Engineering and Design

Evaluation and Simulation-Based Optimization of Urban Traffic Resilience

Qinglong Lu

Vollständiger Abdruck der von der TUM School of Engineering and Design der Technischen Universität München zur Erlangung eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz:

Prof. Dr.-Ing. Rolf Moeckel

Prüfer der Dissertation:

1. Prof. Dr. Constantinos Antoniou
2. Prof. Dr. Jan-Dirk Schmöcker
3. Prof. Dr. Nikolaos Geroliminis

Die Dissertation wurde am 28.03.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Engineering and Design am 17.07.2024 angenommen.

Abstract

Transportation system disruptions often result in significant social and economic consequences, particularly with the increasing frequency of triggers like natural disasters and man-made incidents. This necessitates effective methods for evaluating and optimizing system performance under disruption conditions. Resilience, a comprehensive measure describing systems' ability to withstand and recover rapidly from disruptions, has become a central focus in this research direction. Traffic resilience, a term created for assessing transportation system resilience from the perspective of traffic flows, is yet to be comprehended at the methodological level and implemented at the practical application level. This dissertation aims to develop methods, algorithms, and experiments to address this problem. A conceptual relation diagram is constructed to describe the interdependence of the essential components for traffic resilience assessment frameworks. Specifically, the diagram illustrates that dynamic travel demand modeling and reliable traffic resilience indicators are two prerequisites for the evaluation and optimization of urban traffic resilience. The dissertation contributes to these three research domains across data sources, modeling frameworks, practical applications, and empirical insights.

In terms of origin-destination (OD) demand estimation, methods relying on conventional data sources like household travel surveys and traffic network detection encounter limitations due to low observability and indeterminateness, particularly during disruptions. The dissertation explores the potential of opportunistic data, particularly location-based social network (LBSN) data, to overcome these limitations. More specifically, we develop a dynamic OD matrix estimation method based on LBSN data by leveraging a two-stage stochastic programming framework, with specific consideration of the unique characteristics of LBSN data regarding confirmed trip purpose identification and activity chain hosting. On the other hand, a three-step framework integrating histogram-based feature engineering, clustering, and classification models is proposed to estimate demand pattern changes under demand-side disruptions using crowdsensed data and station characteristics. These methods provide a way to capture demand changes and accurately estimate disrupted travel demand by solely utilizing publicly available data.

Further, we propose novel traffic resilience indicators according to the Macroscopic Fundamental Diagram (MFD) dynamics before and after disruptions, given that the MFD is an intrinsic property of a homogeneously congested transportation network. More importantly, we distinguish the indicators for different types of disruptions, including demand-side (e.g., hyper-congestion), supply-side (e.g., link closure), and mixed disruptions, due to the discrepancy between the system's manifestations in MFD dynamics when encountering these disruptions. Inspired by the serving rate of facilities defined in the queueing theory and the primary role of transportation in facilitating the movement of people and goods, resilience loss is measured by the reduction in trip

completion rate. Furthermore, we explore the relationship between resilience loss and network topological attributes, such as centrality and connectivity, from a variety of synthetic disruption experiments to evaluate the extent to which network topology can explain traffic resilience.

To address simulation-based optimization for urban traffic resilience, the dissertation investigates the feasibility of using surrogate models as a time-efficient alternative to resource-intensive simulations. A physics-informed surrogate modeling framework incorporating the proposed MFD-based traffic resilience indicators is developed, which combines a physical component derived from an analytical network model with a functional component for modeling the difference between simulation results and approximations. Then, we apply this framework to post-disruption recovery measure optimization problems for examination and validation purposes, in which lane reversal control is employed as the recovery measure. Decision-makers can utilize this framework to quickly design appropriate recovery measures in response to disruptive events.

The proposed models, frameworks, and algorithms are evaluated through experiments conducted in large-scale real-world networks, including Tokyo, Kyoto, and Munich.

Zusammenfassung

Verkehrsunterbrechungen führen häufig zu erheblichen sozialen und wirtschaftlichen Folgen, insbesondere aufgrund der zunehmenden Häufigkeit von Auslösern wie Naturkatastrophen und von Menschen verursachten Vorfällen. Dies erfordert wirksame Methoden zur Bewertung und Optimierung der Systemleistung unter Störungsbedingungen. Resilienz, ein umfassendes Maß für die Systemleistung unter Störungen, ist zu einem zentralen Schwerpunkt in dieser Forschungsrichtung geworden. Verkehrsresilienz, ein Begriff, der zur Bewertung der Resilienz von Verkehrssystemen aus der Perspektive von Verkehrsströmen geschaffen wurde, muss noch auf methodischer Ebene verstanden und auf praktischer Anwendungsebene implementiert werden. Diese Dissertation zielt darauf ab, Methoden, Algorithmen und Experimente zu entwickeln, um dieses Problem anzugehen. Ein grundlegendes Beziehungsdiagramm wird erstellt, um die Interdependenz der notwendigen Komponenten für einen Rahmen zur Bewertung der Verkehrsresilienz zu beschreiben. Insbesondere veranschaulicht das Diagramm, dass die dynamische Schätzung der Reisenachfrage und zuverlässige Indikatoren für die Verkehrsresilienz zwei Voraussetzungen für die Bewertung und Optimierung der städtischen Verkehrsresilienz sind.

Im Bereich der Schätzung von Ursprungs-Ziel (OD) Nachfragen stoßen herkömmliche Methoden, die sich auf Datenquellen wie Haushaltsreisebefragungen und Verkehrsnetzerkennung verlassen, aufgrund geringer Beobachtbarkeit und Unbestimmtheit, insbesondere während Störungen, auf Grenzen. Diese Dissertation erforscht das Potenzial von opportunistischen Daten, insbesondere von datenbasierten sozialen Netzwerken (LBSN), um dieses Problem zu lösen. Genauer gesagt, wird mithilfe eines zweistufigen stochastischen Programmierungsrahmens eine dynamische OD-Matrix-Schätzmethode unter Verwendung von LBSN-Daten entwickelt, wobei deren Vorteile bei der Bestätigung des Reisezwecks und Hosting der Aktivitätenketten berücksichtigt werden. Andererseits wird ein dreistufiger methodischer Rahmen vorgeschlagen, der histogrammbasierte Merkmalskonstruktion, Clusterbildung und Klassifizierungsmodelle integriert, um Änderungen im Nachfragemuster des öffentlichen Verkehrs (ÖV) aufgrund von nachfrageseitigen Störungen unter Verwendung von crowdsensing-Daten und Stationsmerkmalen zu schätzen.

Bezüglich der Indikatoren für die Verkehrsresilienz werden neuartige Indikatoren vorgeschlagen, die aus der Dynamik des makroskopischen Fundamental Diagramms (MFD) vor und nach Störungen abgeleitet sind. Da MFD eine intrinsische Eigenschaft eines homogen verstopften Verkehrsnetzes ist und es Unterschiede in den Einflussmechanismen gibt, die verschiedenen Arten von Störungen zugrunde liegen, messen diese Indikatoren separat die Resilienzverluste durch Stau und Netzwerkunterbrechungen. Inspiriert von der Bedienungsrate von Einrichtungen, wie sie in der Warteschlangentheorie definiert ist, und der primären Funktion des Verkehrs, Menschen und Güter zu befördern, wird der Resilienzverlust durch die Verringerung der Reiseabschlussrate gemessen. Darüber hin-

aus untersuchen wir die Beziehung zwischen Resilienzverlust und netzwerk-topologischen Attributen wie Zentralität und Konnektivität anhand einer Vielzahl von synthetischen Störungsexperimenten, um zu bewerten, inwieweit die Netzwerk-Topologie die Verkehrsresilienz erklären kann.

Um die simulationsbasierte Optimierung der städtischen Verkehrsresilienz anzugehen, untersucht diese Dissertation die Machbarkeit der Verwendung von Surrogatmodellen als zeit-effiziente Alternative zu ressourcenintensiven Simulationen. Es wird ein physikinformierter Surrogatmodellierungsrahmen entwickelt, der die vorgeschlagenen auf MFD basierenden Verkehrsresilienzindikatoren integriert und eine physische Komponente aus einem analytischen Netzwerkmodell mit einer funktionalen Komponente kombiniert, um die Unterschiede zwischen Simulationsergebnissen und Annäherungen zu modellieren. Anschließend wenden wir diesen Rahmen auf das Optimierungsproblem von Maßnahmen zur Wiederherstellung nach Störungen an, bei dem die Umkehrung der Fahrspur als Wiederherstellungsmaßnahme eingesetzt wird. Entscheidungsträger können diesen Rahmen nutzen, um schnell geeignete Wiederherstellungsmaßnahmen als Reaktion auf Störungsereignisse zu entwerfen.

Die vorgeschlagenen Modelle, Rahmen und Algorithmen werden durch Experimente in großen realen Netzwerken, einschließlich Tokyo, Kyoto und München, bewertet.

Acknowledgements

"To dust is gone the fame achieved in thirty years; Like cloud-veiled moon the thousand-mile plain disappears". These two lines translated from an ancient Chinese poem, aptly capture my emotions at this moment. For nearly seven years, I have nurtured the dream of reaching this milestone, and it is with profound gratitude that I acknowledge the support and guidance of those who have accompanied me on this journey.

Foremost, I extend my deepest gratitude to my supervisor, Prof. Constantinos Antoniou, for granting me this opportunity to pursue my aspirations. His wealth of experience and keen insight into research have been instrumental in guiding me through critical junctures, enabling the progression of my doctoral project within a structured framework while contributing meaningfully to the field of study. His constructive feedback during our discussions has been invaluable, facilitating the smooth execution of my research endeavors. Moreover, his mentorship in career planning has instilled in me the confidence to navigate my academic trajectory.

I am also indebted to Dr. Moeid Qurashi, Dr. Wenzhe Sun, Prof. Jan-Dirk Schmöcker, and Dr. Vishal Mahajan, Cheng, and Jiannan for their contributions to my research. Their expertise in their respective domains has enriched our discussions and collaborations, leading to fruitful outcomes. I am grateful for the opportunities to co-author publications with them. To Dr. Mohamed Abouelela, for his constructive comments on the structure of my dissertation. I am grateful to my colleagues and visiting scholars in the laboratory, for their support and cooperation, which has fostered a conducive and enjoyable work environment. Special thanks to Cheng, Hao, and Danyue for the memorable moments we shared during our dining and outings.

Heartfelt thanks are due to my family and friends for their unwavering support and encouragement throughout this journey. Their belief in me has been a constant source of strength, reassuring me that I am not alone in this endeavor.

Lastly, I express my gratitude to the Bundesministerium für Bildung und Forschung (BMBF) and the Deutscher Akademischer Austauschdienst (DAAD) for their financial support.

Contents

List of Figures

List of Tables

Acronyms

ABM	Activity-Based Model.
AFC	Automated Fare Collection.
AHA	Adaptive Hyperbox Algorithm.
AIC	Akaike Information Criterion.
APC	Automated Passenger Counting.
API	Application Programming Interface.
BDS	BeiDou Navigation Satellite System.
BIC	Bayesian Information Criterion.
COVID-19	Corona Virus Disease 2019.
DODE	Dynamic Origin-Destination Demand Estimation.
DTA	Dynamic Traffic Assignment.
DUE	Dynamic User Equilibrium.
EN	Elastic Net.
FFPT	Fare-Free Public Transport.
FHWA	Federal Highway Administration.
GA	Genetic Algorithm.
GBD	Generalized Benders Decomposition.
GD	Gradient Descent.
GLS	Generalized Least Squares.
GMM	Gaussian Mixture Model.
GoF	Goodness-of-Fit.
GPR	Gaussian Process Regression.
GPS	Global Positioning System.
GPT	Google Popular Times.
LBSN	Location-Based Social Network.
MC	Monte Carlo.
MCMC	Markov chain Monte Carlo.
MFD	Macroscopic Fundamental Diagram.

MILP	Mixed-Integer Linear Programming.
OD	Origin-Destination.
PCA	Principal Component Analysis.
POI	Point of Interest.
PT	Public Transport.
PTDM	Public Transport Demand Management.
RNDP	Road Network Design Problem.
RO	Research Objective.
RQ	Research Question.
SAA	Sample Average Approximation.
SPSA	Simultaneous Perturbation Stochastic Approximation.
SUMO	Simulation of Urban MObility.
TAZ	Traffic Analysis Zone.
TSSP	Two-Stage Stochastic Programming.
UE	User Equilibrium.
UNDRR	United Nations Office for Disaster Risk Reduction.

Chapter 1

Introduction

1.1 Background and motivation

Natural disasters and special events, whether planned or unplanned, can easily result in disruptions to urban transportation systems. These disruptions have demonstrated their devastating effects on people's daily lives, as well as social and economic activities, as can be inferred from the ramifications of their triggers. For instance, in 2023, 28 natural disaster events caused more than a billion dollars in damage per event in the United States alone (NOAA, 2024). Among these events, two resulted in 247 and 100 deaths, respectively, while an additional five events caused more than 10 deaths each.

Additionally, the frequency of natural disasters has notably increased in the past two decades. According to a report collated by the United Nations Office for Disaster Risk Reduction (UNDRR), the number of disaster events between 2000 and 2019 was 1.74 times that in the previous twenty years (between 1980 and 1999), resulting in approximately $2.97 trillion in economic losses. The disruptions caused by these disasters, for instance, road inundation due to flooding (increases about 2.34 times within the same period), impose a significant challenge on transportation system operations. Therefore, it is imperative to develop a comprehensive methodological framework for establishing reliable assessment and enhancement mechanisms to assist in improving the performance of transportation systems in the face of disruptive events.

The term *resilience* originates from the Latin word "resiliere", which means bounce back or spring back. It was first conceptualized and applied in academic research by Holling (1973) specifically in evaluating ecological systems. Since then, it has been introduced to various disciplines, including organization, economics, social science, engineering, etc. (Zhou et al., 2019). In the evaluation of system performance under disruptions, resilience is deemed more comprehensive compared with other terms proposed for similar purposes, such as vulnerability and robustness. This is because it explicitly accounts for the additional recovery phase, including the new equilibrium state and recovery speed, which is, however, usually overlooked by other terms. Mathematically, system resilience can be quantified by integrating the deviation of system functionality from its optimal state throughout the disruption period (Bruneau et al., 2003). This can be expressed as Equation (1.1), where t_0 and t_1 indicate the start and end of the disruption, respectively, and $Q(t)$ represents the system functionality at time t. For ease of understanding, Figure 1.1 provides a graphical illustration of the calculation of resilience loss, commonly referred to as the "resilience triangle" in the literature. Upon the occurrence of a disrup-

tion, the system experiences a decline in functionality. However, owing to the system's inherent resilience and the implementation of recovery measures, this decline eventually stabilizes and initiates a process of restoration towards a new equilibrium state. Hence, the entire disruption period can often be divided into a disruption phase and a recovery phase.

$$R = \int_{t_0}^{t_1} (100 - Q(t))\,\mathrm{d}t \tag{1.1}$$

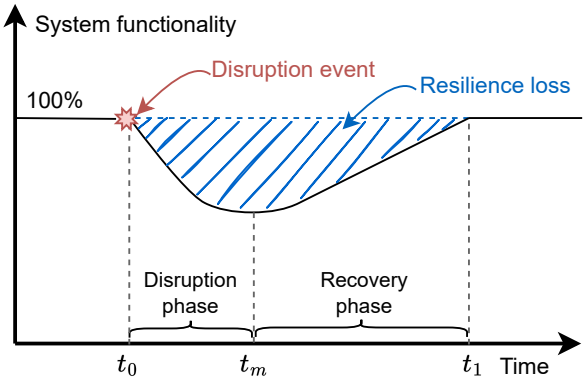


Figure 1.1: Graphical definition of system resilience.

The application of the resilience concept to transportation systems has attracted much attention from scholars and practitioners in recent years (see reviews from Zhou et al., 2019; Gonçalves and Ribeiro, 2020; Pan et al., 2021; Serdar et al., 2022). Whilst transportation system resilience is defined slightly differently from one another across relevant publications to accommodate specific application requirements, a typical definition given by the Federal Highway Administration (FHWA) of the United States is as follows: "Resilience represents the ability to prepare for changing conditions and withstand, respond to, and recover rapidly from disruptions". Moreover, it is worth noting that disruptions in transportation systems often come with recurring or non-recurring congestion, thereby impeding traffic efficiency. Accordingly, *traffic resilience* has been proposed to evaluate the traffic states over the course of a given disruption. Here, traffic resilience is defined as "the ability of an urban road transportation system to prepare for different kinds of disruptions, effectively serve vehicles, and recover rapidly to its optimal serving rate" (Lu et al., 2024c). Compared to conventional indicators for evaluating transportation system resilience, reliable traffic resilience indicators offer a reasonable alternative in traffic-oriented evaluations of transportation system performance under disruptions.

The Origin-Destination (OD) matrix, as a representative of mobility demand patterns, is indispensable in urban transportation management and traffic control systems (Xiong et al., 2020; Qurashi et al., 2022; Lu et al., 2024b). Combining it with a traffic assignment model enables the investigation of urban traffic flows in a detailed manner. Hence, the OD matrix usually serves as a fundamental component of numerous traffic operational

applications. Clearly, accurate travel demand estimates are also crucial to the reliability of traffic resilience evaluation. To date, conventional OD estimation methods typically rely on either household travel surveys or traffic measurements collected by fixed loop detectors spread throughout the network and floating cars (e.g., taxis). However, the former category of methods is often limited to long-term urban planning applications owing to the low frequency of survey implementation (Jin et al., 2014), while the latter structurally suffers from the indeterminateness issue due to the insufficiency of detectors compared to the high-dimensional OD flows (Cascetta et al., 2013; Antoniou et al., 2016; Qurashi et al., 2022). Furthermore, in the context of disruptive events, these conventional methods are prone to encountering missing data issues resulting from potential infrastructure damage. These limitations make them inappropriate for dynamic estimation of disrupted travel demand. Mahajan (2023) suggested that streaming data from emerging data sources can help address this challenge. Emerging data here refers to data types emerging from relatively newer sources, for instance, mobile crowdsensed social media (Harrison et al., 2020). It has also been highlighted that these data can provide disrupted travel demand management with reliable and up-to-date information (Pender et al., 2014). As a result, leveraging these opportunistic data for travel demand estimation provides opportunities to overcome the limitations inherent in conventional methods in the context of transportation disruptions.

On the other hand, a reliable indicator of system functionality is also critical to the evaluation of system resilience. As per Yang et al. (2023), about 50 indicators have been proposed to quantify transportation system resilience. These indicators can be primarily classified as network topology-based indicators (e.g., link criticality) and trip-based indicators (e.g., average travel time). While network topology-based indicators are capable of providing insight into the structural risk of transportation networks, and trip-based indicators can reflect traffic efficiency to some extent, they often fail to incorporate information on transportation network characteristics, traffic dynamics, and travel demand patterns simultaneously. Therefore, they may not fully represent transportation system resilience as a holistic property of the system.

Existing resilience optimization approaches for transportation systems can be mainly categorized into big data-driven, simulation-based, mathematical programming, and probabilistic methods (Serdar et al., 2022). Among these, simulation-based methods hold the superiority of integrating disaggregate models of stochastic driver behaviors (e.g., car-following, lane-changing, route choice), which, once calibrated using real data, can provide accurate and detailed traffic state estimates for a given strategy (Osorio and Chong, 2015). However, traffic simulation models are generally computationally expensive due to their complex, dynamic, and stochastic nature, restricting their applications in many practical problems that require fast decisions, such as emergency recovery measure optimization. Consequently, it is beneficial to develop efficient models to cope with this issue so as to improve the use of simulation-based approaches in practice. Surrogate modeling approaches, which combines the advantages of traffic simulators in estimating time-dependent traffic dynamics and the efficiency of mathematical optimization, have been proposed for this purpose.

To summarize, with accurate dynamic demand estimates and reliable resilience indicators, one is able to create a testbed to examine various resilience enhancement strategies. The optimization of urban traffic resilience thus becomes possible. We present the conceptual relation diagram in Figure 1.2 to illustrate this essential argument. Simply speaking, it says that demand modeling and resilience indicators are two prerequisites for traffic resilience evaluation and optimization.

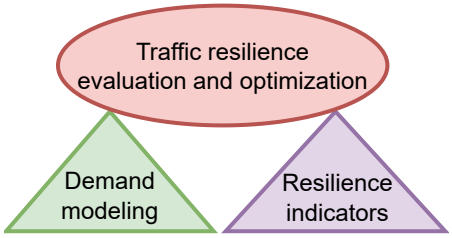


Figure 1.2: Relation diagram of the three essential components in traffic resilience assessment frameworks.

1.2 Problem definition and research objectives

Given the huge social and economic losses resulting from disruptions to transportation systems, analyzing their response to and recovery from such disruptions is always of great interest to the transportation community. Although the resilience concept has been extensively employed to assess the performance of urban transportation systems under disruptions, traffic resilience, emphasizing the fundamental functions of transportation, i.e., faciliatating the movement of people and goods, is a relatively new term in the literature. A comprehensive traffic resilience assessment framework relies on three essential components: dynamic travel demand modeling, traffic resilience indicators, and traffic resilience optimization approaches. To date, however, the state-of-the-art has yet to provide satisfactory and convincing solutions in these three directions. It follows that the Research Questions (RQs) listed below remain open for researchers, which are also the focus of the dissertation.

RQ(1) How can we precisely estimate the travel demand under disruptions?

RQ(2) What indicators can be employed to reliably measure traffic resilience in urban transportation systems?

RQ(3) How can we accurately formulate and efficiently solve traffic resilience optimization problems?

Following these RQs, this dissertation aims to advance the modeling of disrupted travel demand, as well as the evaluation and optimization of urban traffic resilience. Given the limitations of conventional data sources in the estimation of travel demand under disruptions and the great potential of emerging opportunistic data in providing

real-time information on urban mobility patterns, it is expected to obtain benefits from developing travel demand estimation methods using opportunistic data. On the other hand, since state-of-the-art microscopic traffic simulation models can produce complex and reliable traffic dynamics via their detailed modeling of the intertwined interactions between system supply and demand, they have become a popular tool for addressing transportation operation and control problems. However, no studies have been conducted to explore their applications to traffic resilience optimization problems to date. With these considerations, the following Research Objectives (ROs) are formulated:

RO(1) Investigate the applications of emerging opportunistic data in travel demand estimation.

RO(2) Develop robust and reliable traffic resilience indicators to evaluate the performance of urban transportation systems under disruptions.

RO(3) Explore the relationship between traffic resilience indicators and conventional resilience indicators adopted in transportation systems.

RO(4) Develop simulation-based approaches for traffic resilience optimization.

RO(5) Develop computationally efficient algorithms to tackle city-wide traffic resilience optimization problems.

1.3 Dissertation contributions

This dissertation contributes to the existing body of knowledge on travel demand modeling, traffic resilience, and simulation-based optimization. In particular, we summarize the main contributions of the dissertation below.

- Travel demand estimation methods built upon emerging opportunistic data.
- Robust traffic resilience indicators based on Macroscopic Fundamental Diagrams (MFDs).
- Regression analysis relating traffic resilience with transportation network topology.
- Surrogate modeling for simulation-based traffic resilience optimization problems.
- Post-disruption recovery measure optimization model for improving traffic resilience.
- Efficient algorithms for solving discrete simulation-based optimization problems.

Detailed contributions are specified in the respective chapters of the dissertation.

1.4 Dissertation structure and list of publications

The structure of the dissertation is shown in Figure 1.3. The dissertation is centered and focused around the central diagram (also provided in Figure 1.2) describing the interrelation among travel demand modeling, traffic resilience indicators, and traffic resilience evaluation and optimization. The current chapter and the subsequent chapter, through reviewing the state-of-the-art literature, identify research gaps, formulate RQs, and delineate ROs. The central conceptual relation diagram (Figure 1.2) is also an outcome of this analysis. The next chapter after that expands the relation diagram to develop a methodological framework for traffic resilience evaluation and optimization. For each component in the diagram, two chapters with different research topics are embedded, contributing to the understanding of the corresponding field from both theoretical and practical perspectives. Finally, the dissertation concludes with a summary of research findings and insights gleaned from the studies, and outlines future research directions.

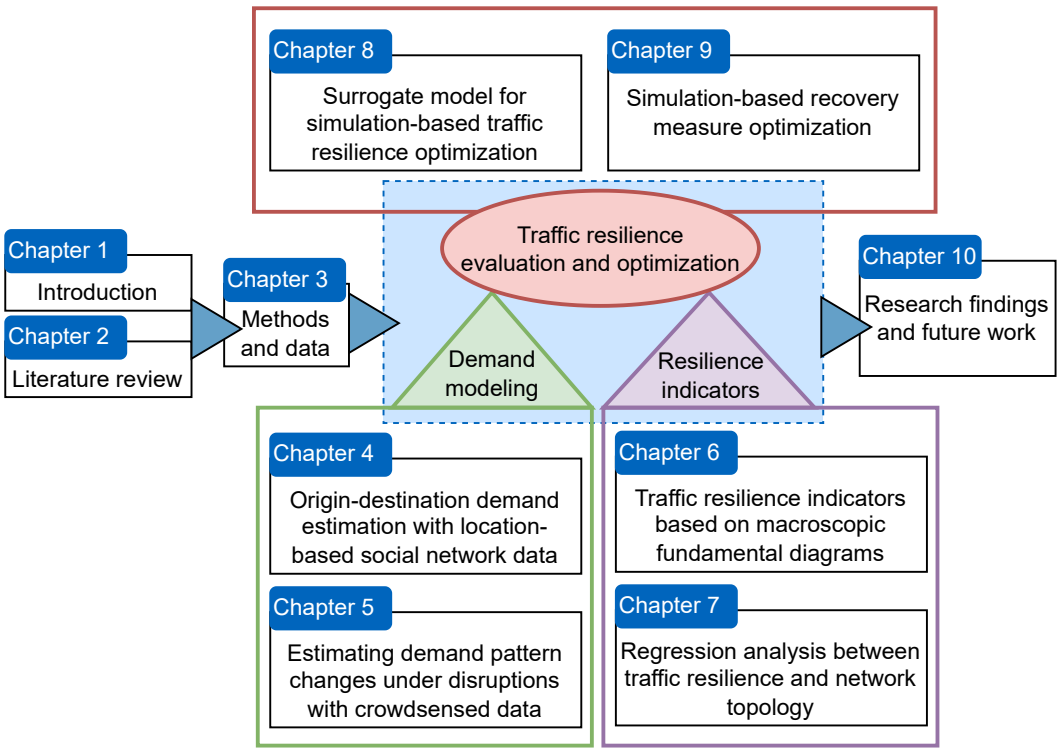


Figure 1.3: Dissertation structure.

We briefly describe the key contents of each chapter as follows.

- Chapter 1 introduces the background, motivation, and research questions, objectives, contributions and the structure of this dissertation. It presents the conceptual relation diagram illustrating the three essential components of a traffic resilience assessment framework and their interdependent relationship, directing the research endeavors of the dissertation.

- Chapter 2 reviews the state-of-the-art literature related to the ROs of this dissertation. It covers topics such as travel demand estimation, indicators and optimization approaches for transportation system resilience, and the simulation-based optimization approaches in transportation engineering. This chapter aims to identify literature gaps crucial for addressing the ROs.

- Chapter 3 presents the overall methodological framework of the dissertation, which is also a general framework for investigating traffic resilience evaluation and optimization problems. It also summarizes the models, solution algorithms, data, and study areas used in the following chapters.

- Chapter 4 develops an OD estimation model driven by location-based social network data leveraging the two-stage stochastic programming framework. We propose an algorithm that integrates Markov chain Monte Carlo sampling and the generalized Benders decomposition algorithm to solve the model.

- Chapter 5 develops a three-step busyness-based framework involving a histogram-based feature engineering approach, a clustering model component, and a classification model component to investigate the potential of crowdsensed data for capturing demand pattern changes in public transport systems under demand disruptions.

- In Chapter 6, inspired by the fact that the Macroscopic Fundamental Diagram (MFD) is an intrinsic property of a homogeneously congested transportation network, we proposes MFD-based traffic resilience indicators by associating the trip completion rate and system functionality. Resilience to demand-side and supply-side disruptions are separately considered due to the distinction in their influencing mechanisms.

- Chapter 7 provides a regression analysis of the relationship between traffic resilience indicators and conventional network topology-based indicators. It illustrates the extent to which structural risks can explain traffic resilience. On the other hand, it can serve as a general guidance for the construction of traffic-oriented resilient transportation networks.

- Chapter 8 develops a physics-informed surrogate model to expedite the solution algorithm for simulation-based traffic resilience optimization problems. The model integrates a dynamic and temporally correlated analytical network model to offer information on the physical laws between the variables of interest, and a functional component to capture the difference between analytical and simulated values.

- Chapter 9 applies the proposed surrogate model to address post-disruption recovery measure optimization problems with an objective to improve traffic resilience. The lane reversal technique is creatively used as recovery measures under supply disruptions.

- Chapter 10 concludes by summarizing the proposed methods, findings, and insight from the experiments conducted, and also suggests possible future research directions.

This dissertation comprises a collection of the following publications resulting from the research endeavors described herein.

Journal articles:

- Lu, Q.-L., Qurashi, M., and Antoniou, C. (2023b). Simulation-based policy analysis: The case of urban speed limits. *Transportation Research Part A: Policy and Practice*, 175:103754.

- Lu, Q.-L., Mahajan, V., Lyu, C., and Antoniou, C. (2024a). Analyzing the impact of fare-free public transport policies on crowding patterns at stations using crowdsensing data. *Transportation Research Part A: Policy and Practice*, 179:103944.

- Lu, Q.-L., Qurashi, M., and Antoniou, C. (2024b). A two-stage stochastic programming approach for dynamic OD estimation using LBSN data. *Transportation Research Part C: Emerging Technologies*, 158:104460.

- Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2024c). Traffic resilience quantification based on macroscopic fundamental diagrams and analysis using topological attributes. *Reliability Engineering & System Safety*, 247:110095.

- Lu, Q.-L., Sun, W., Lyu, C., Schmöcker, J.-D., and Antoniou, C. (2024d). Post-disruption lane reversal optimization with surrogate model to improve urban traffic resilience. *(Under Review)*.

Peer-reviewed conferences:

- Lu, Q.-L., Qurashi, M., and Antoniou, C. (2022). A stochastic programming method for od estimation using lbsn check-in data. In *4th Symposium on Management of Future Motorway and Urban Traffic Systems (MFTS 2022)*.

- Lu, Q.-L., Qurashi, M., and Antoniou, C. (2023c). A two-stage stochastic programming approach for dynamic OD estimation. In *102nd Annual Meeting of the Transportation Research Board (TRB 2023)*.

- Lu, Q.-L., Mahajan, V., Lyu, C., and Antoniou, C. (2023a). Analyzing crowding in public transport during 9-EUR ticket using crowdsensing data. In *25th Euro Working Group on Transportation Meeting (EWGT 2023)*.

– Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2023d). An MFD-based optimization approach to improve transportation system resilience under infrastructure disruption. In *25th Euro Working Group on Transportation Meeting (EWGT 2023)*.

– Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2023e). Surrogate modeling for recovery measure optimization to improve traffic resilience. In *9th International Symposium on Transport Network Resilience (INSTR 2023)*.

Chapter 2

Literature review

The content of this chapter has been partially presented in the following works:

Lu, Q.-L., Mahajan, V., Lyu, C., and Antoniou, C. (2024a). Analyzing the impact of fare-free public transport policies on crowding patterns at stations using crowdsensing data. *Transportation Research Part A: Policy and Practice*, 179:103944.

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2024b). A two-stage stochastic programming approach for dynamic OD estimation using LBSN data. *Transportation Research Part C: Emerging Technologies*, 158:104460.

Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2024c). Traffic resilience quantification based on macroscopic fundamental diagrams and analysis using topological attributes. *Reliability Engineering & System Safety*, 247:110095.

Lu, Q.-L., Sun, W., Lyu, C., Schmöcker, J.-D., and Antoniou, C. (2024d). Post-disruption lane reversal optimization with surrogate model to improve urban traffic resilience. *(Under Review)*.

2.1 Introduction

This chapter reviews the literature and the state-of-the-art approaches and applications in the field pertaining to the three essential components of traffic resilience assessment: travel demand modeling, resilience indicators, and resilience evaluation and optimization. As such, this chapter is structured into three sections, each dedicated to one specific aspect.

In the first section, we delve into the issue of scalability inherent in conventional traffic-measurement-based OD demand estimators, which stands as a major limitation of these methods. We review the endeavors that focus on eliminating this limitation, which becomes particularly pronounced when considering the conditions under disruption situation. Given the potential of opportunistic data in either replacing, complementing or supplementing conventional data in OD estimation (Pender et al., 2014; Mahajan et al., 2023), we also review the state-of-the-art in using opportunistic data, in particular Location-Based Social Network (LBSN) data and crowdsensed data, for demand modeling.

In the second section, we conduct a comprehensive review of network-wide resilience indicators proposed for transportation systems. In particular, according to the re-

search scope of this dissertation, we primarily focus on two categories: those based on simulation-based approaches and those developed using the concept of MFD.

In the third section, we review the literature on major disruption categories that have been broadly investigated in the transportation community and the evaluation approaches adopted. Additionally, we are also interested in approaches to recovery measure optimization aimed at improving transportation system resilience. Recognizing the computational cost associated with traffic simulation models, we then summarize surrogate modeling approaches to transportation simulation-based optimization problems to improve their computational efficiency.

According to the findings obtained from the review, we have identified the research gaps for each research topic, which are detailed in the respective sections of this chapter.

2.2 Dynamic travel demand modeling

2.2.1 Traffic-measurement-based estimators

State-of-the-art traffic measurement systems, such as loop detectors, capture the effect of the mobility demand on the network and not the demand itself (Frederix et al., 2011; Shafiei et al., 2017). Therefore, traffic-measurement-based dynamic OD estimators structurally suffer from the issue of indeterminateness in estimating realistic OD flow patterns and estimate rather the fluctuations with respect to an existing prior OD estimate from survey/planning models (McNally, 2007) using time-dependent traffic measurements, instead of the whole demand itself. Furthermore, such methods require mapping the OD estimates into a comparable set of traffic measurements to formulate a traffic-measurement-based objective function. Therefore, an evident distinction in such models is the use of an assignment matrix, where assignment matrix-based algorithms explicitly use an analytical representation of the relationship between demand and traffic counts for Dynamic Origin-Destination Demand Estimation (DODE) (Cascetta and Postorino, 2001; Toledo and Kolechkina, 2012), which is usually assumed to be linear and is not the case in reality, especially for large-scale and complex networks. Thus, recent years saw a shift towards assignment matrix-free methods which allow more accurate modeling of the supply and demand relationship using Dynamic Traffic Assignment (DTA) and can also incorporate other data sources which can not be analytically related to OD flows (e.g., Bluetooth data). However, even the popular assignment matrix-free algorithms, e.g., Simultaneous Perturbation Stochastic Approximation (SPSA) (Balakrishna et al., 2007), fail to estimate large-scale problems, since the gradient approximation gets highly sensitive against increasing problem non-linearity, while the computational requirements (i.e., the number of iterations and running DTA iteratively) also increase exponentially.

To address the scalability issues of DODE, recent literature moved towards either reducing the OD dimensions or reducing problem non-linearity by adding structural/-correlation information in the objective function. The most prominent scalability approach is the use of Principal Component Analysis (PCA) (Djukic et al., 2012; Qurashi et al., 2019, 2022) that directly reduces OD dimensions and non-linearity transforming OD matrix into lower dimensional orthogonal Principal Components using the extracted

variance in historical OD estimates. Although powerful and intuitive, PCA-based methods, strongly rely on the presence and quality of historical estimates to extrapolate estimation patterns. Next comes the use of quasi-dynamic methods that use a correlation assumption to reduce the DODE estimation variables (Cascetta et al., 2013; Cantelmo et al., 2014; Bauer et al., 2017). At first, Cascetta et al. (2013) proposed the assumption of keeping OD shares constant with varying generation profiles of each origin, significantly reducing the estimation variables. While later to eliminate the requirement for a historical OD matrix in the original framework, Bauer et al. (2017) combined a Generalized Least Squares (GLS) estimator for link flows and a maximum entropy term for the non-observable traffic distribution across paths in the modified framework, assuming instead of constant OD shares, constant path choice proportions over time-of-day intervals for days with similar mobility patterns. Besides them, two other prominent SPSA scalability approaches also exist that explicitly attempt to reduce problem non-linearity via either using weight matrices to add OD flows and network correlation information in the DODE objective function (Antoniou et al., 2015) or clustering homogeneous model parameter (Tympakianaki et al., 2015).

Although all the above-mentioned efforts do significantly improve DODE scalability, they still fundamentally require setting up dynamic traffic models alongside the estimation algorithm setup and running them iteratively (varying by their convergence efficiency). Furthermore, almost all such conventional approaches apart from Cantelmo et al. (2020) do not consider modeling activity chains and trip purpose information during estimation. However, OD flows are an aggregated representation of individuals' activity-travel chains.

2.2.2 Origin-destination demand estimation based on location-based social network data

Leveraging the benefits of high resolution in spatial and temporal dimensions at a low cost, LBSN check-in data has emerged as a reliable secondary data source for travel demand estimation (Hu and Jin, 2017; Silva et al., 2019). The unique feature of self-contained confirmed trip purposes (activities) has given LBSN an unparalleled advantage in OD estimation. LBSN data, a kind of geotagged social media data, is capable of estimating distance and duration distributions, OD matrices, and individual activity-based mobility patterns (Zhou et al., 2018). Furthermore, it is also a supplementary and reliable data source for analyzing urban activities and behaviors (Rizwan et al., 2020). Consequently, some LBSN data-based approaches have been proposed to understand and estimate OD demand.

Previous studies have shown the feasibility of using LBSN check-in data for demand pattern estimation. For instance, Hu and Jin (2015) applies a simple time-dependent model to estimate trip attraction and validates the feasibility of using LBSN check-in data in demand pattern estimation. Due to the heterogeneity of venue categories in demand patterns, they suggested that trip arrival patterns of different venue categories should be considered separately. Similarly, Yang et al. (2014b) and Yang et al. (2015) apply an integrated model that combines a hierarchical clustering method for venue cat-

egorization and a regression model to estimate trip production and attraction. Trips are then distributed according to a singly constrained gravity model. However, these models require calibration with the ground truth OD data provided by the local urban planning agency. Cebelak (2013) and Jin et al. (2014) further improve this approach by replacing the singly constrained gravity model with a doubly constrained one. The experiment results show that the modified method can significantly reduce the OD estimation error caused by the sampling bias from the singly constrained model. Additionally, other conventional trip distribution models for urban travel demand analysis, such as the radiation model, rank-based model, and population-weighted opportunities model, with inputs extracted from LBSN data, have also been tested and compared in Kheiri et al. (2015). However, these models can only provide a static (day-level) solution to the OD estimation problem.

Inspired by the promising performance of the Hawkes process in self-reinforcing behavior modeling in Cho et al. (2014), Hu and Jin (2017) presents a time-of-day zonal arrival estimation model by integrating the Hawkes process and an LBSN check-in observation model into a state-space modeling framework. This approach can reduce the sampling bias in OD estimation caused by the difference between social behaviors and real travel patterns. In addition, Hu et al. (2019) incorporates a Latent Dirichlet Allocation (LDA) model for profiling zonal functionality based on the venue categorical distribution and a Pearson Product-Moment Correlation (PPMC) analysis method for measuring pairwise zone correlation. These two methods can help improve the performance of the zonal trip arrival and OD estimation models in the previous model, which is crucial for the performance of LBSN-data-based OD estimators (Jin et al., 2014).

However, none of the existing literature has fully utilized the trip purpose and activity chain information contained in LBSN data. Moreover, most models were constructed based on conventional gravity models, which require additional information about the population and network. Additionally, they overlooked the uncertainty factor of user social behaviors.

2.2.3 Demand pattern analysis with crowdsensed data

The proliferation of smartphones equipped with positioning technologies has provided a novel means of collecting mobility data, opening up new avenues for Public Transport (PT) policy evaluation in a more cost-effective manner (Tse et al., 2018). These devices generate vast volumes of real-time data on individuals' activities and mobility through location-based services, social networks, and other mobile applications. Some smartphone datasets have been leveraged for urban mobility analysis for purposes such as PT network design (e.g., Pinelli et al., 2016, using mobile phone trajectories) and PT accessibility analysis (e.g., Cai et al., 2017, using mobile phone signaling data). Unlike these normally unavailable sensitive personal data, crowdsourced data, such as geotagged tweets (Chaniotakis et al., 2017), provide new opportunities to gain more insights into urban mobility, including demand patterns, without increasing privacy exposure risks (Capponi et al., 2019; Vitello et al., 2023). However, despite the rapid growth of crowd-

sourced data, their potential for analyzing demand patterns, particularly in the context of disruptions, has not been fully explored (Niu and Silva, 2020).

Crowdsensed check-in rates or busyness data at Point of Interests (POIs) is a typical kind of crowdsourced data. The raw location data from mobile devices near POIs are collected and transmitted via an internet connection to a central server. These raw data from numerous personal devices are processed, anonymized, and aggregated to provide the busyness data at POIs. Among others, the real-time busyness data of PT stations, representing the crowding patterns, contains valuable information about PT demand patterns. Therefore, one can evaluate the impact of disruptions by a before-after comparison of the crowding patterns extracted from busyness data. These emerging crowdsensed data have wide coverage and fine resolution, enabling a detailed empirical analysis of the impact of disruptions on PT demand patterns at a large scale.

Examples of such crowdsensed data sources include Foursquare check-ins (D'Silva et al., 2018) and popularity trends (Capponi et al., 2019; Timokhin et al., 2020). Previous studies have successfully extracted spatiotemporal demand patterns and activities from these data (Timokhin et al., 2020; MacKenzie and Cho, 2020; Capponi et al., 2019; Möhring et al., 2021). Live POI check-in rates have also been used to analyze the impact of special events or interventions, such as lockdowns during the pandemic of Corona Virus Disease 2019 (COVID-19), on POI demand patterns (Mahajan et al., 2021), the impact of natural disasters, such as heavy snowfalls, on people daily activities within the urban area (Santiago-Iglesias et al., 2023).

2.3 Transportation system resilience indicators

2.3.1 Resilience indicators in simulation-based evaluation

The evaluation of transportation system resilience has traditionally focused on purely topological measures derived from the theory of complex network analysis, such as connectivity (e.g., alpha index) and accessibility metrics (e.g., betweenness centrality) (see Table 1 in Zhang et al., 2015; Pan et al., 2021, for a review). While these static indicators provide some insight into the structural risk of transportation networks, they fail to characterize the impact of traffic demand changes and the development and cascading effects of disruptions in the spatiotemporal dimension (Shekar et al., 2017; Zeng et al., 2019; Xu and Chopra, 2022). In addition, these indicators also fail to evaluate the influence of a partial loss of capacity due to sub-links failure (there are usually multiple sub-links between two nodes in transportation networks) (Guidotti et al., 2017). Since disruptions often come with traffic congestion, to comprehensively assess the resilience of transportation networks, it is essential to take into consideration traffic dynamics that capture the time-varying nature of network congestion.

In a comparative analysis conducted by Balal et al. (2019), multiple traffic dynamics-oriented resilience measures were examined, including queue length, link speed, link travel time, frontage road delay, and detour route delay. The study found a relatively low correlation among these measures, suggesting that the importance of links to network resilience depends on the specific measure employed. This finding implies that

evaluating network resilience at the link level may lead to contradictory conclusions and offer limited contributions. Consequently, the need to reliably measure network performance in both spatial and temporal dimensions has pushed the evaluation of resilience beyond local traffic performance (e.g., travel time, speed, and delays of individual links or intersections) towards the development of network-wide measures (Miller-Hooks et al., 2012; Gao et al., 2016).

Bucar and Hayeri (2020) presented a framework for assessing the effects of precipitation-induced flood events in urban areas using a two-layer simulation model. Metrics such as vehicle miles traveled, vehicle hours traveled, and trips completed were employed to evaluate system resilience. Similarly, Hoogendoorn et al. (2015) evaluated network resilience by measuring changes in level-of-service due to density deviation caused by incidents. Average speed served as a proxy for level-of-service, making system resilience a function of density and density deviation. However, these indicators overlook the comparison between disruption and normal operation situations, as well as the time required for recovery, which are important considerations in defining resilience. Additionally, they make the calculation of the optimal system functionality (i.e., the equilibrium state before disruption) challenging, as they heavily depend on demand levels and patterns.

On the other hand, Chen and Miller-Hooks (2012) and Zhang et al. (2015) utilized the ratio of the maximum post-disaster throughput to the pre-disaster throughput as a resilience indicator. Chen and Miller-Hooks (2012) employed this measure to optimize post-event recovery actions, while Zhang et al. (2015) used it to assess the significance of network topology in resilience. Yao and Chen (2023) applied the percolation theory, using the giant component size (i.e., the size of the largest connected component of the network), to evaluate network resilience to random failures. However, these indicators overlook traffic dynamics and their time-varying nature, thus providing an inaccurate depiction of resilience along the temporal dimension.

2.3.2 Resilience indicators based on macroscopic fundamental diagrams

In recent years, there has been significant interest among scholars in resilience indicators based on the theory of MFD, since it represents an intrinsic property of a homogeneously congested transportation network. One of the pioneering efforts in this direction in the field of transportation resilience was made by Kim and Yeo (2017), who proposed a network performance loss indicator utilizing the concept of MFD to assess the criticality of network links. The indicator takes into account that disruptive events lead to increased spatial heterogeneity in traffic flow, thereby altering the shape of the MFD. The performance loss is quantified as the proportion of the aggregated flow reduction caused by disruptions, given by

$$FR = \frac{\int_{k_0}^{k_c} q(k) - q^d(k)\ \mathrm{d}k}{\int_{k_0}^{k_c} q(k)\ \mathrm{d}k} \tag{2.1}$$

where $q(k)$ and $q^d(k)$ are the weighted space-mean flow across the network under normal and disruption operation conditions, respectively; k_0 and k_c represent the lower and upper bounds for comparison. However, in their definition, both $q(k)$ and $q^d(k)$ are

weighted based on the total length of the original network, without accounting for link removal during disruptions. This weighting scheme can lead to inaccurate comparisons (Amini et al., 2018). Furthermore, the comparison solely focuses on the shape of the MFD, disregarding the temporal dynamics of the network. Consequently, the comparison may introduce biases in cases where the MFD shape remains unchanged despite the occurrence of disruptive events.

To address these limitations, Amini et al. (2018) introduced an improved indicator by considering trip length changes in the evaluation of network resilience. They defined the service rate of a network as the number of completed trips per unit of time. Building upon the linear relationship discovered by Geroliminis and Daganzo (2008) between trip production (P) and trip completion rate (D), i.e., $P = \mathcal{L}D$, where $\mathcal{L}$ represents the average trip length, and leveraging the relationship between trip production and weighted flow, i.e., $P = \sum_{i \in \mathbb{L}} q_i l_i = qL$, where L denotes the length of all links equipped with detectors, they derived a resilience index that reflects the network's service rate using weighted flow as a surrogate. The index is calculated as follows:

$$RI(t) = q(t) - q^d(t) \times \frac{\mathcal{L}}{\mathcal{L}_d} \tag{2.2}$$

where $\mathcal{L}$ and $\mathcal{L}_d$ are the average trip length under normal and disruption conditions, respectively. We note that this index has been normalized by $L/\mathcal{L}$ without distinguishing the network before and after disruptions. It means, the same as in Kim and Yeo (2017), the length of the network is assumed unchanged during disruptions, which violates the reality of infrastructure disruptions.

Further, Gao et al. (2022) introduced a definition of traffic resilience for a system comprising two reservoirs, utilizing the total congestion deviation, which is measured as the difference between the operational vehicle accumulation and the optimal value. Assuming parabola-shaped MFDs, the traffic resilience is computed through the following equation:

$$TR = -\int_{t_0}^{t_c} \left| n_1(t) - \frac{n_1^{\max}}{2} \right| + \left| n_2(t) - \frac{n_2^{\max}}{2} \right| \, \mathrm{d}t \tag{2.3}$$

where $n_i(t)$ is the vehicle accumulation at reservoir i at time t, and $n_i^{\max}$ is the corresponding maximum accumulation The time interval from t_0 to t_c represents the congestion period. A larger congestion deviation (the term inside the integral at a specific t) represents severer congestion or more spare capacity. Note that this indicator is designed for evaluating the resilience to hyper-congestion. However, characterizing capacity spare as a measure of resilience loss may be unreasonable since it is primarily determined by limited travel demand rather than network malfunction. Resilience, as an indicator of system functionality, should not be affected by capacity spare. Additionally, using accumulation as a metric for transportation system performance, particularly during disruptions, has limitations. Firstly, the number of vehicles within the network, represented by accumulation, does not provide a comprehensive reflection of transportation network functionality. The performance/functionality of a system/facility is generally represented by its service rate (i.e., the trip completion rate in MFD) rather than the

number of customers (i.e., the accumulation in MFD). Secondly, the optimal accumulation during disruptions can exceed or equal that of normal operations, implying that the same accumulation value may result in lower resilience loss during disruptions, leading to inaccurate estimations.

2.4 Transportation system resilience optimization

2.4.1 Transportation system disruptions and resilience evaluation

Seismic resilience and flood resilience are primary focuses due to the severe consequences of earthquakes and floods. Studies on seismic resilience primarily focus on post-earthquake recovery optimization. Given the role that bridges play in the resilience of road–bridge transportation networks, scheduling the recovery of bridges has been extensively studied. Various metrics have been utilized to determine bridge recovery schedules, including network connectivity and vehicular traffic (Ahmed et al., 2022), recovery time and cumulative benefit rate (Zhang et al., 2022a), network travel time (Zhang et al., 2022b), and recovery time and skewness of the recovery curve (Somy et al., 2022). There are also some studies only performed pure assessment of resilience to seismic hazards, such as Virtucio et al. (2024), and Zhang et al. (2023c). Regarding flood resilience, various flood scenarios have been investigated, such as hurricane-induced flood (Dong et al., 2023), and precipitation-induced flood (Bucar and Hayeri, 2020; Wassmer et al., 2024). These investigations often rely on empirical analyses using historical data focusing on the influence on network topology, reliability, and stability (Dong et al., 2023; Bucar and Hayeri, 2020; Wassmer et al., 2024). Noteworthy, Zhang and Alipour (2023) proposed a two-stage stochastic programming approach to optimize pre-disruption mitigation and post-disruption recovery measures simultaneously to enhance resilience to flood hazards, where data-driven simulation was employed. In addition to seismic resilience and flood resilience, snow storm resilience and typhoon resilience have also received increasing attention in recent years, such as Fang et al. (2022), Mirjalili et al. (2023), and Santiago-Iglesias et al. (2023).

Some studies also explored resilience equality issues. For instance, Byun and D'Ayala (2022) provided a probabilistic analysis of disruption inequality across urban areas regarding seismic resilience. The inequality problem of disruptions has also been examined from the perspectives of socioeconomic groups and access to amenities (Wei et al., 2022; Anderson et al., 2022).

On the other hand, evaluating the resilience of multi-modal transportation systems has emerged as a research hotspot in recent years, given the interdependence among different urban transportation systems such as subway, bus, and taxi. Examples include Wang et al. (2023) (bus, subway and taxi), Chen et al. (2023) (bus and metro), Zhang et al. (2023b) (focusing on node reliability), Aparicio et al. (2022) (focusing on topological robustness), and Fang et al. (2022) (focusing on complementary transport modes). These studies collectively contribute to a deeper understanding of the resilience dynamics inherent in multi-modal transportation systems, thus providing valuable insights for enhancing their overall resilience in the face of disruptions.

Various models have been applied to address resilience-based evaluation and optimization problems, with notable categories including probabilistic models, mathematical optimization models, and data-driven models. Probabilistic models usually integrate a hazard component with a probability distribution of disturbance intensity, an evaluation component quantifying resilience, and a simulation component for mapping strategies to resilience outputs. For example, Amini and Padgett (2023) proposed a methodology integrating hazard and debris exposure models to assess the impact of hurricane-induced debris on transportation network resilience. Similar frameworks were presented in Taghizadeh et al. (2023) and Wu and Chen (2023) to evaluate the seismic resilience of transportation systems during emergency medical responses. Additionally, Zhang et al. (2023a) applied a probabilistic recovery model to measure the restoration of bridges after earthquakes. In contrast, mathematical optimization models are applied to assist in planning pre-disruption mitigation and post-disruption recovery strategies, subject to practical constraints such as budget, computational time, and personnel constraints (Serdar et al., 2022). Representative models include two-stage stochastic programming (e.g., Zhang and Alipour, 2023), and bi-level optimization (e.g., Amghar et al., 2024). However, these models often suffer from computational inefficiency, particularly in large-scale transportation networks.

On the other hand, data-driven methods leverage sensor data, such as loop detector data and Global Positioning System (GPS) data, collected over an extended period. With advancements in sensor technology and database storage and processing, these methods have also gained significant attention. Moreover, they hold promise for providing real-time information on resilience. For instance, Diab and Shalaby (2020) used metro system data to evaluate the impact of outdoor tracks and weather conditions on system resilience. Roy et al. (2019) proposed a method to detect disruptions from geo-located social media data and assess their impact on mobility resilience. A hybrid knowledge-based and data-driven approach was developed in Yin et al. (2022) with the aid of the Bayesian Network model and historical metro data to identify component weakness and enhance resilience with proper improvement. For a comprehensive review and comparison of these model categories, including their merits and limitations, readers are referred to Serdar et al. (2022).

In summary, while the aforementioned studies vary in terms of the disruptions of interest and the methods used for solution, they all rely on appropriate resilience indicators for evaluation and optimization. Thus, there is a pressing need for a reliable resilience indicator to facilitate the application of the resilience concept in urban road transportation systems from the perspective of traffic flow.

2.4.2 Recovery measure optimization for transportation system resilience

Determining economically viable post-disruption recovery measures is critical to resilience assessment and optimization (Xu et al., 2022). Notably, recovery measure optimization for road–bridge transportation networks, among others, has attracted substantial attention, given the pivotal role of bridges in the resilience of such networks. Different scenarios, ranging from urban (e.g., Zhang and Alipour, 2020) to rural envi-

ronments (e.g., Zhang et al., 2022a), necessitate distinct methodologies and resilience emphases. Urban scenarios often emphasize social-economic outcomes, while rural contexts prioritize minimizing total recovery time and time delay. In contrast, Zhang et al. (2017) utilized a synthetic network for evaluation, proposing a recovery action scheduling optimization for road–bridge networks by minimizing total recovery time and the skew of the recovery curve.

Studies have also focused on resilience optimization for systems that couple transportation systems with other correlated functional systems. For instance, Ge et al. (2019) developed a Mixed-Integer Linear Programming (MILP) model to optimize the post-storm recovery process (repair scheduling) in power and transportation systems simultaneously, given their heavy interdependence. Similarly, in optimizing the resilience of power distribution systems, Li et al. (2022) integrated the optimization of repair crew allocation on the road network into a two-stage modeling framework to explicitly measure the influence of their interdependence. In addition, Pei et al. (2024) formulated a two-stage recovery strategy optimization model for the interdependent transportation and healthcare system post-earthquake, to maximize seismic resilience.

Rather than exclusively focusing on recovery measures, some studies embed optimizations for both preparedness and recovery measures into an integrated modeling framework. Stochastic programming, owing to its capability to incorporate a set of disruption scenarios, is a commonly used technique, such as in the road–bridge network resilience optimization problem (Zhang and Alipour, 2020), the OD connectivity (a resilience indicator) optimization problem (Liao et al., 2018), and the network topology evaluation (w.r.t. network resilience) problem (Zhang et al., 2015). However, these methods often build on analytical formulations with strong simplification assumptions (e.g., recovery measures are defined as capacity increases), making them less practical and easier to implement in real applications. As a result, the case studies analyzed therein are on either synthetic networks (e.g., in Ge et al., 2019; Li et al., 2022) or simple, small real networks (e.g., in Chen and Miller-Hooks, 2012; Zhang and Alipour, 2020) tractable by their models.

To more accurately approximate real situations and provide a realistic evaluation of system functionality, scholars resort to adding more components representing traffic states in their modeling processes. Examples include considering user travel behaviors (Zou and Chen, 2021), employing traffic assignment methods (Pei et al., 2024), incorporating uncertainty quantification via sampling algorithms (Xiao et al., 2022; Nan and Sansavini, 2017; Zhang and Alipour, 2020), modeling user equilibrium (Zhao and Zhang, 2020; Li et al., 2019), and using simulation-based objective evaluation (Liao et al., 2018). However, these detailed considerations significantly increase computational costs due to the exploration of large solution spaces and the expensive functionality evaluation (Zou and Chen, 2021). This runs counter to the rapidity property of system resilience (Bruneau et al., 2003; Pei et al., 2024), which requires timely, efficient, and effective recovery measures after disruptive events (Liu et al., 2021; Aydin et al., 2018; Alkhaleel et al., 2022; Wang and Wang, 2020). Therefore, although simulation-based modeling approaches provide a more accurate way to quantify resilience (Ahmed and Dey, 2020), efficient solution algorithms are crucial to increase their values in practical

applications, especially for large-scale transportation networks. However, this aspect has been addressed in relatively few works.

2.4.3 Surrogate modeling for simulation-based urban road transportation optimization

Simulation models for time-dependent traffic dynamics at the individual level are usually computationally expensive, requiring several minutes to over an hour for a single run, depending on the network scale (Chen et al., 2014). Large-scale transportation optimization problems involving high-dimensional decision variables and uncertainty would further increase the computational cost of simulation-based objective functions. Consequently, it easily suffers from the difficulty of extensively exploring the decision space within a limited computational budget (López and Monzón, 2010). Metaheuristic algorithms are commonly adopted to address large-scale problems due to their superiority in handling high-dimensional optimization (Osorio and Chong, 2015; Chen et al., 2015). However, treating the simulator as a black box without considering structural relationships among variables of the underlying transportation problem can lead to intensive recalls of objective evaluations to identify the direction of performance improvement (Osorio and Chong, 2015; Chen et al., 2015). Large-scale simulations, thus, are usually limited to what-if scenario analysis.

Surrogate model (also known as metamodel) based simulation optimization is one feasible approach to address such high-dimensional optimization problems with expensive-to-evaluate objective functions. Surrogate modeling is a macro-modeling technique that attempts to replicate the response surface describing the relationship between decision variables and simulation outputs (Forrester et al., 2008). In other words, the surrogate model replaces the complex traffic simulator with a simplified analytical approximation, enabling faster computation. This approach combines the advantages of simulation in estimating time-dependent traffic dynamics and the efficiency of mathematical optimization.

Surrogate modeling has found applications in various simulation-based transportation optimization problems. Examples include identifying faulty behavior regions of autonomous vehicles (Beglerovic et al., 2017), optimizing highway toll charges (Chen et al., 2014), and freeway dynamic OD estimation (Huo et al., 2023). While conventional general-purpose (functional) surrogate models mainly include low-order polynomials, spline models, radial basis function, and Gaussian Process Regression (GPR), state-of-the-art surrogate models integrate an additional physical component that yields structural information about the problem to expedite algorithm convergence. Osorio and Bierlaire (2013) combined an analytical queueing network model (Osorio and Bierlaire, 2009) with a quadratic polynomial to address traffic signal control problems for a small network of 12 links. The computational efficiency of the initial queueing network model was later improved in Osorio and Chong (2015) and Chong and Osorio (2018) to enable city-scale traffic signal control. This physics-informed surrogate model methodology was then applied to high-dimensional offline time-independent OD calibration in Osorio (2019b) and time-dependent OD calibration in Osorio (2019a), where

the analytical queueing network model was replaced with a network model integrating a simple route choice model, a flow propagation model, and a speed-density relationship model, to fit the characteristics of OD calibration problems. Furthermore, Zhou et al. (2023) extended this approach to discrete simulation-based optimization for (two-way) car-sharing network design problems (i.e., fleet allocation across stations to maximize the expected profit).

Despite its advantages in computational efficiency and promising performance uncovered in the literature, few efforts have been made to optimize post-disruption recovery measures using surrogate modeling. Although artificial neural network-based surrogate models have been adopted to assess seismic risk in road-bridge transportation networks (Yoon et al., 2020) and optimize post-earthquake recovery measures for interdependent transportation and healthcare systems (Pei et al., 2024), they relied on less accurate mathematical models, providing no time-dependent information. No studies have proposed simulation-based recovery measure optimization via surrogate models.

Chapter 3

Methods and data

The content of this chapter has been partially presented in the following works:

Lu, Q.-L., Mahajan, V., Lyu, C., and Antoniou, C. (2024a). Analyzing the impact of fare-free public transport policies on crowding patterns at stations using crowdsensing data. *Transportation Research Part A: Policy and Practice*, 179:103944.

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2024b). A two-stage stochastic programming approach for dynamic OD estimation using LBSN data. *Transportation Research Part C: Emerging Technologies*, 158:104460.

Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2024c). Traffic resilience quantification based on macroscopic fundamental diagrams and analysis using topological attributes. *Reliability Engineering & System Safety*, 247:110095.

Lu, Q.-L., Sun, W., Lyu, C., Schmöcker, J.-D., and Antoniou, C. (2024d). Post-disruption lane reversal optimization with surrogate model to improve urban traffic resilience. *(Under Review)*.

3.1 Introduction

This chapter introduces the methodological framework for traffic resilience evaluation and optimization, followed by an introduction to the fundamental methods for the topics involved in the framework, and data and study areas for experiments conducted in subsequent chapters.

The chapter is structured as follows: The next section presents the overall methodological framework for implementing traffic resilience evaluation and optimization. Subsequently, we introduce the fundamental models and algorithms applied to solve the problems highlighted in the framework in Section 3.3. Finally, we present the data used and study areas considered in the dissertation in Section 3.4.

3.2 Methodological framework

In Chapter 1, we presented a conceptual relation diagram illustrating the three essential components—demand modeling, resilience indicators, and optimization approaches—in traffic resilience evaluation and optimization, as depicted in Figure 1.2. In this section, we expand upon that diagram by incorporating specific modules for each component,

as illustrated in Figure 3.1. The expanded diagram serves as the methodological framework of the dissertation, which, in fact, is also a general framework for traffic resilience research. In the following, we provide a detailed introduction to the modules and their interactions displayed in the framework sequentially, respecting the order of components mentioned above.

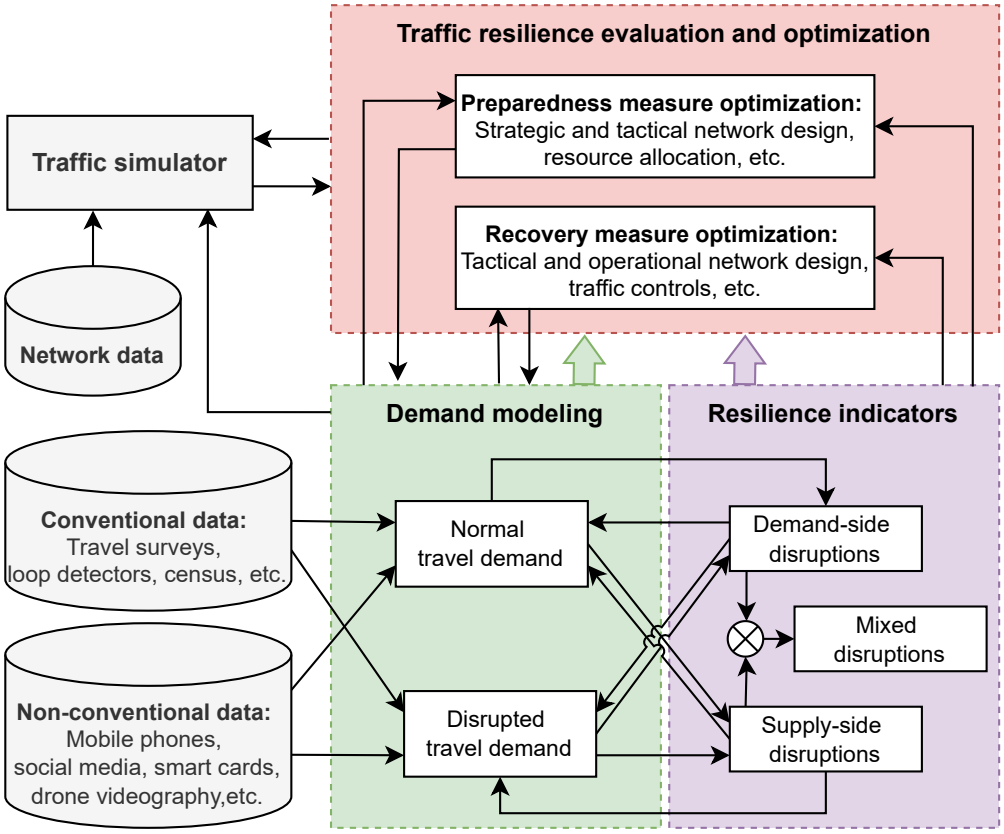


Figure 3.1: Overall methodological framework for traffic resilience evaluation and optimization.

3.2.1 Disrupted travel demand modeling

In the demand modeling component of the methodological framework (Figure 3.1), we add two modules to account for the alteration of travel demand under different operational conditions of transportation systems. To be specific, normal and disrupted travel demand estimation are the main focuses in this component.

Data sources for transportation analysis and modeling can generally be categorized into conventional and non-conventional (emerging) data, as outlined by Mahajan (2023). Figure 3.1 enumerates several types of data that have been commonly used in travel demand modeling. However, collecting conventional data tends to be time-consuming and expensive, and their declining availability and usability during transportation disruptions highlight the need for demand estimation models built upon emerging data sources.

The dissertation focuses on estimating disrupted travel demand. There are two ways to accomplish this task: directly estimating the disrupted demand or estimating demand pattern changes, from which disrupted demand can be derived by subtracting the changes from normal travel demand, as depicted in Figure 3.2. Both conventional and emerging data sources are valuable in either approach. However, since conventional data are prone to data missing issues during disruptions and emerging data offer advantages in terms of availability and coverage, we primarily focus on developing approaches based on emerging data sources. Specifically, opportunistic data has emerged as a promising candidate to complement, supplement, or replace conventional data in travel demand modeling, as discussed in Chapters 1 and 2. Therefore, in Chapter 4, we develop an LBSN-based OD estimator, and in Chapter 5, we explore the potential of crowdsensed data for estimating demand pattern changes.

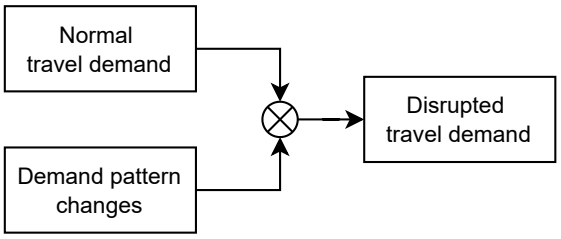


Figure 3.2: Disrupted travel demand modeling.

3.2.2 Traffic resilience indicators

Transportation system disruptions can be categorized into demand-side disruptions, supply-side disruptions, and mixed disruptions, depending on their impact on the demand and supply sides of the system. Mixed disruptions are characterized by the presence of both demand-side and supply-side disruptions. The empirical analysis, illustrated in Chapter 6, suggests that the resilience indicators for different categories of disruptions should be distinguished owing to the distinct manifestations of the system. In Chapter 6, we address this need by developing different indicators for measuring traffic resilience to different categories of disruptions based on MFD dynamics. Furthermore, Chapter 7 aims to deepen the understanding of traffic resilience by conducting a regression analysis on its relationship with previous transportation resilience indicators, such as network topological attributes.

Nevertheless, it is worth mentioning that all disruptions are intertwined with travel demand due to the inherent interdependence between system supply and demand, which is also indicated in Figure 3.1, emphasizing the complex relationship between disruptions and travel demand across the transportation system.

3.2.3 Traffic resilience evaluation and optimization

After preparing disrupted travel demand and resilience indicators, traffic resilience evaluation and optimization can be performed. These optimization problems can be classified

into two main categories: preparedness measure optimization and recovery measure optimization, according to the existing literature on transportation resilience optimization. Resilience indicators typically serve as the objective or part of the objective function in these optimization problems, although they may also only serve as evaluation metrics for optimization results.

Most transportation resilience optimization problems can be considered as network design problems due to their similarity in the modeling process. According to Farahani et al. (2013), network design problems can be categorized into strategic, tactical, and operational design problems based on the nature of decision variables in terms of their time span and impact on infrastructures. Preparedness measure optimization problems mainly encompass strategic and tactical network design and resource allocation. Recovery measure optimization problems mainly include tactical and operational network design and traffic control problems. Note that tactical network design can be either preparedness or recovery optimization, depending on the scenario characteristics and decision variables.

In general, preparedness measures intend to enhance the system's ability to withstand disruptions, while recovery measures attempt to improve the system's recovery during disruptions. Both preparedness and recovery measures interact with travel demand. However, whilst preparedness measures are more likely to affect normal travel demand in the long run, recovery measures have an immediate impact on disrupted demand. We emphasize that the focus of the dissertation is on simulation-based optimization in order to leverage the advantages of advanced traffic simulators in modeling complex interactions among system demand, supply, and travelers. Network data and travel demand are essential inputs to the traffic simulator. In Chapter 8, a surrogate model is constructed to expedite the algorithm for solving simulation optimization problems without losing the benefits of simulations. Subsequently, Chapter 9 utilizes the surrogate model to address recovery measure optimization problems in the immediate aftermath of disruptions.

3.3 Methods

This section summarizes some fundamental methods that are used to support the modeling and solution development for the methodology presented in the next chapters. Table 3.1 shows the chapters where these methods are used.

3.3.1 Optimization algorithms

3.3.1.1 Sample average approximation

In Chapter 4, we develop an LBSN-based OD estimator by utilizing the Two-Stage Stochastic Programming (TSSP) framework. In Chapter 8, we establish a surrogate-based simulation optimization model. Both models consider the uncertainty in travel and driving behaviors. The resulting problems are non-convex and difficult to get solved as the expectation $\mathbb{E}_\xi$ embedded in the objective function is usually an integral of a

Table 3.1: Summary of methods.

Method	Chapter
Optimization algorithms	
Sample average approximation	Chapter 4, Chapter 8
Generalized Benders decomposition algorithm	Chapter 4
Machine learning techniques	
Markov chain Monte Carlo sampling	Chapter 4
LightGBM classification model	Chapter 5
Simulation-based optimization fundamentals	
Generic formulation of simulation-based optimization	Chapter 8, Chapter 9
Surrogate model for traffic simulators	Chapter 8, Chapter 9

complex function. ξ is a random variable describing the problem state. Accordingly, in practice, we often need to assume ξ has a finite number of possible realizations, such that we can estimate $\mathbb{E}_\xi$ by

$$\mathbb{E}_\xi \left[f(\mathbf{x}, \xi) \right] = \sum_{n}^{N} p_n F(\mathbf{x}, \xi_n) \tag{3.1}$$

where $\mathbf{x}$ is the decision variable, N is the total number of realizations, p_n denotes the probability of realization n, and f denotes the stochastic objective function. However, the probability distribution of these realizations is often unavailable. To address this, applying an appropriate sampling technique and the Sample Average Approximation (SAA) method, the expectation can be approximated as

$$\mathbb{E}_\xi \left[f(\mathbf{x}, \xi) \right] \approx \frac{1}{N_s} \sum_{n}^{N_s} F(\mathbf{x}, \xi_n) \tag{3.2}$$

where N_s is the total number of scenario samples.

3.3.1.2 Generalized Benders decomposition algorithm

The TSSP problem for estimating OD matrices presented in Chapter 4 can be solved via decomposition algorithms. We apply the Generalized Benders Decomposition (GBD) algorithm, which was first proposed in Geoffrion (1972) for addressing the mathematical programming problems with complicating variables (i.e., variables that if fixed to given values render a simple or decomposable problem). In TSSP, the first-stage decision variables are the complicating variables of the problem. The idea behind GBD is to decompose the original problem into a master problem and a series of subproblems (one per scenario). In the master problem, the first-stage decisions, denoted by $\mathbf{x}$, are optimized. In the subproblems, the second-stage decisions, denoted by $\mathbf{y}$, are optimized. They are solved alternately until convergence. At a specific iteration k, the subproblems

are solved first separately resulting in the optimum $\mathbf{y}^k(\xi)$ given $\mathbf{x}^{k-1}$ and scenario ξ. An optimality cut (or feasibility cut) is constructed based on the dual solutions of subproblems (or feasibility problem), which is added to the master problem as a new constraint. Given all cut constraints created through a pass-back mechanism from subproblems in previous iterations, the master problem is solved with respect to $\mathbf{x}$ resulting in $\mathbf{x}^k$. Note, these cuts gradually reduce the feasible space of the complicating variable.

To describe the algorithm, we sequentially provide the formulations of the subproblem (SP), the feasibility problem (FP), and the master problem (MP). At the k-th iteration, for a given scenario ξ_n and $\mathbf{x}^{k-1}$, the SP is formulated as follows:

$$\begin{aligned}
(\text{SP}) \min_{\mathbf{y}} \quad & f_2(\mathbf{y}, \xi_n) && (3.3)\\
\text{s.t.} \quad & \text{Constraints in the second stage} && (3.4)\\
& \mathbf{x} = \mathbf{x}^{k-1} : \quad \lambda_n^k && (3.5)
\end{aligned}$$

where $f_2(\cdot)$ denotes the objective function of the second-stage problem. Compared to the original second-stage problem, SP has an additional equality constraint, Equation (3.5), which fixes the complicating variable $\mathbf{x}$ to the optimal value from the previous iteration. The solution of SP provides values to $\mathbf{y}^k$ in different scenarios, as well as the corresponding optimal Lagrange multipliers vector associated with Equation (3.5), i.e., the optimal dual variables, λ^k. However, it is possible that not all SPs are solvable, whilst the feasibility of SPs is directly relative to the construction of Bender's cut at the respective algorithm iteration, i.e., optimality cut or feasibility cut. If SP is feasible, the optimality cut would be established using the terms in the Lagrangian function of SP that are relevant to $\mathbf{x}$. The formulation is given below.

$$\mathcal{L}_o(\mathbf{x}, \mathbf{y}^k(\xi_n), \lambda_n^k) = f_2(\mathbf{y}^k(\xi_n), \xi_n) + (\lambda_n^k)^\top (\mathbf{x} - \mathbf{x}^{k-1}) \tag{3.6}$$

However, if SP is infeasible, the following FP needs to be solved.

$$\begin{aligned}
(\text{FP}) \min_{\mathbf{y},\eta} \quad & \eta && (3.7)\\
\text{s.t.} \quad & \text{Constraints in the second stage} && (3.8)\\
& \eta \geq 0 && (3.9)\\
& \mathbf{x} - \mathbf{x}^{k-1} \leq \eta : \quad \mu_n^k && (3.10)
\end{aligned}$$

Similarly, we can get the Lagrangian multiplier vector μ_n^k for the constraints expressed by Equation (3.10). The Lagrangian function of FP is given by (the terms irrelevant to $\mathbf{x}$ have been eliminated)

$$\mathcal{L}_f(\mathbf{x}, \mu_n^k) = (\mu_n^k)^\top (\mathbf{x} - \mathbf{x}^{k-1} - \eta) \tag{3.11}$$

In this case, a feasibility cut is created based on the Lagrangian function of FP. Once SPs and/or FPs have been tackled, MP will be updated by adding new constraints of Benders cut in the current iteration. MP is formulated as

$$\text{(MP)} \min_{\mathbf{x},\alpha} \quad f_1(\mathbf{x}) + \alpha \tag{3.12}$$

$$\text{s.t.} \quad \text{Constraints in the first stage} \tag{3.13}$$

$$\frac{w_s}{N_s} \sum_{n=1}^{N_s} \mathcal{L}_o(\mathbf{x}, \mathbf{y}^t(\xi_n), \lambda_n^t) \leq \alpha \qquad \forall t \in \mathbb{I}_o \tag{3.14}$$

$$\mathcal{L}_f(\mathbf{x}, \mu_l^t) \leq 0 \qquad \forall l \in \mathbb{S}_f^t, \forall t \in \mathbb{I}_f \tag{3.15}$$

where $f_1(\cdot)$ denotes the objective function of the first-stage problem, $\mathbb{I}_o$ is the set of indices of the iterations at which all SPs are feasible, $\mathbb{I}_f$ is the set of indices of the iterations at which at least one of the SPs is infeasible, and $\mathbb{S}_f^t$ is the set of scenarios whose associated SPs are infeasible at iteration t. Equations (3.14) are denominated as optimality cuts, while Equations (3.15) are feasibility cuts. From the MP, we can get the values of the first-stage decision variables $\mathbf{x}^k$.

Algorithm 1 Generalized Benders decomposition algorithm.

1: Initialize the first-stage variable $\mathbf{x}_0$.
2: Initialize the iteration index $k = 1$, complicating variables $\mathbf{x}^k = \mathbf{x}_0$, error tolerance ϵ, maximum number of iterations M.
3: Set the lower bound of the objective function $\underline{z}^k = 0$, and the upper bound $\overline{z}^k = \infty$.
4: **while** $|\overline{z}^k - \underline{z}^k|/|\underline{z}^k| \geq \epsilon$ and $k < M$ **do**
5: Set $k := k + 1$.
6: Solve the subproblems by fixing $\mathbf{x}$ as $\mathbf{x}^{k-1}$.
7: **if** all subproblems are feasible **then**
8: Obtain solution $\mathbf{y}^k$ and the dual variables of those constraints that fix the complicating variables to given values λ^k.
9: Calculate:
$z = f_1(\mathbf{x}^{k-1}) + \frac{w_2}{N_s} f_2(\mathbf{y}, \xi_n)$.
10: Update the upper bound $\overline{z}^k = \min\{\overline{z}^{k-1}, z\}$.
11: Set $\mathbb{I}_o := \mathbb{I}_o \cup \{k\}$.
12: **else**
13: Solve the feasibility problems associated with the infeasible subproblems.
14: Obtain solution $\mathbf{y}^k$, dual variable μ^k and the set of infeasible subproblems $\mathbb{S}_f^k$.
15: Set $\mathbb{I}_f := \mathbb{I}_f \cup \{k\}$.
16: **end if**
17: Add the new optimality cut (or feasibility cuts) to the master problem.
18: Solve the master problem to get $\mathbf{x}^k$ and α^k.
19: Update the lower bound $\underline{z}^k = f_1(\mathbf{x}^k) + \alpha^k$.
20: **end while**

It is worth mentioning that if the original objective function is convex on the complicating variable, GBD can guarantee the strong optimality condition, i.e., the optimal solution from the decomposed problems is equivalent to the original problem. For

convenience, further details on the procedure of the GBD algorithm are presented in Algorithm 1, where N_s denotes the number of scenarios, and w_s denotes the weight of the second-stage objective in the objective function of the first-stage problem.

3.3.2 Machine learning techniques

3.3.2.1 Markov chain Monte Carlo sampling

In Chapter 4, we integrate the Markov chain Monte Carlo (MCMC) sampling method, which has been popularly adopted for scenario generation in stochastic programming, with the SAA method to solve the TSSP. Note, the integration of Monte Carlo (MC) and SAA can reduce the computational costs in generating cuts relative to the discrete-then-solve approach. GBD is one kind of such approach. The modified Metropolis-Hastings algorithm proposed in Au and Beck (2001) is implemented.

The modified Metropolis-Hastings algorithm adopts an accept-reject principle. In the nth step, we denote the current state by ξ_n (high-dimensional) and denote the proposed state by ζ_n. For a specific variable v in the state vector, denote the proposal distribution of the value of v at time τ by $\mathcal{S}_v(\cdot|q_v)$. The proposed state $q_v(\zeta_n)$ is then sampled from $\mathcal{S}_v(\cdot|q_v(\xi_n))$, i.e., $q_v(\zeta_n) \sim \mathcal{S}_v(\cdot|q_v(\xi_n))$. The acceptance probability can be calculated as

$$a_v(q_v(\xi_n), q_v(\zeta_n)) = \min\left\{1,\ \frac{\pi_v(q_v(\zeta_n))\mathcal{S}_v(q_v(\xi_n)|q_v(\zeta_n))}{\pi_v(q_v(\xi_n))\mathcal{S}_v(q_v(\zeta_n)|q_v(\xi_n))}\right\} \tag{3.16}$$

where $\pi_v(\cdot)$ is the known distribution of v at time τ. Finally, the proposed state ζ_n is accepted or rejected based on the principle below.

$$q_v(\xi_{n+1}) = \begin{cases} q_v(\zeta_n) & \text{if } u_v \le a_v(q_v(\xi_n), q_v(\zeta_n)) \\ q_v(\xi_n) & \text{otherwise} \end{cases} \tag{3.17}$$

where u_v is a scalar drawn from a standard uniform distribution, i.e., $u_v \sim \mathcal{U}(0,1)$.

The full description of MCMC sampling is given in Algorithm 2, where $\mathbb{V}$ represents the set of variables in the state vector.

3.3.2.2 LightGBM classification model

In Chapter 5, we apply LightGBM to perform PT station classification based on busyness-based features and station characteristics. LightGBM is an improved version of the gradient boosting decision tree model that has gained popularity in recent years due to its exceptional performance. Traditional decision tree models classify data into different categories using "if-else" decision rules that partition the feature space into subspaces. However, a single decision tree can overfit easily and doesn't generalize well to new data. Gradient boosting is an ensemble technique that improves decision tree performance by training multiple trees sequentially, with each tree trained to correct the errors of previous trees (Bishop, 2006). LightGBM further improves gradient boosting by introducing gradient-based one-side sampling and exclusive feature bundling techniques (Ke et al.,

Algorithm 2 Modified metropolis-Hastings algorithm.

Initialize the number of samples to generate N_s.
Initialize the current state $\xi_0 = \xi$.
Initialize $n = 0$.
while $n < N_s$ **do**
  for each $v = 1, 2, \cdots, |\mathbb{V}|$ **do**
    Simulate $q_v(\zeta_n) \sim \mathcal{S}_v(\cdot|q_v(\xi_n))$.
    Compute the acceptance probability with Equation (3.16)
    Accept or reject with Equation (3.17)
  end for
  Construct the $(n+1)$th state:

$$\mathbf{q}^{(l)}(\xi_{n+1}) = \mathbf{q}^{(l)}(\xi_0)$$
$$\mathbf{q}(\xi_{n+1}) = \left[q_1(\xi_{n+1}), q_2(\xi_{n+1}), \cdots, q_{|\mathbb{V}|}(\xi_{n+1})\right]^\top$$

  Set $n := n + 1$.
end while

2017). The former technique selects data points with the largest gradients during training, focusing on the most difficult samples. Exclusive feature bundling groups relate features together to reduce dimensionality and, thus, improve model efficiency. These unique techniques have made LightGBM one of the most competitive machine learning models for a wide range of applications in both academia and industry (Shwartz-Ziv and Armon, 2022; Bojer and Meldgaard, 2021).

3.3.3 Simulation-based optimization fundamentals

3.3.3.1 Generic formulation of simulation-based optimization

Due to the stochasticity in transportation systems, simulation-based optimization generally estimates the corresponding transportation problems with SAA as

$$\text{(SO)} \quad \min_{\mathbf{x}\in\Omega_{\mathbf{x}}} \quad \hat{f}(\mathbf{x};\mathbf{\Psi}) = \frac{1}{N_s}\sum_{i=1}^{N_s} F_i(\mathbf{x};\mathbf{\Psi}) \tag{3.18}$$

where $\mathbf{x}$ denotes the vector of decision variables, $\Omega_{\mathbf{x}}$ indicates the feasible space, $\mathbf{\Psi}$ indicates the exogenous variables (e.g., network topology) in simulations, N_s is the number of simulation runs, $F_i(\cdot)$ is the i-th realization of the stochastic system performance measure F. $\hat{f}$ can then be regarded an approximation to F using the mean of N_s realizations.

Stochastic traffic simulation can provide considerable information about the effects of implementing specific control decisions. While simulation-based optimization approaches have been widely used to address traffic control and management problems such as what-if analysis in urban planning, practical real-time control problems for alleviating the

negative impact of disruptions often demand quick responses, imposing a challenge on the determination of effective measures, especially for large-scale networks. Surrogate modeling approaches thus have been proposed to address the real-time simulation-based optimization problems.

3.3.3.2 Surrogate model for traffic simulators

Surrogate models have been proposed to accelerate the solution for simulation-based optimization problems, such as Osorio and Bierlaire (2013); Chong and Osorio (2018); Osorio (2019a). Making use of a similar structure to the ones developed in these studies, we propose a surrogate model that integrates a physical (problem-specific) component and a functional (general-purpose) component. The proposed surrogate model has the form of

$$\mathscr{M}(\mathbf{x}, \mathbf{z}; \boldsymbol{\beta}, \boldsymbol{\theta}) = \eta f^A(\mathbf{x}, \mathbf{z}; \boldsymbol{\theta}) + \phi(\mathbf{x}; \boldsymbol{\beta}) \tag{3.19}$$

where $f^A(\cdot)$ represents the physical component, and $\phi(\cdot)$ represents the functional component. Generally, the physical component is derived from a analytical model, while the functional component is calculated from a regression model. $\mathbf{z}$ denotes the endogenous variables in the physical component, while $\boldsymbol{\theta}$ denotes the exogenous variables. $\boldsymbol{\beta}$ denotes the parameters to be estimated in the functional component.

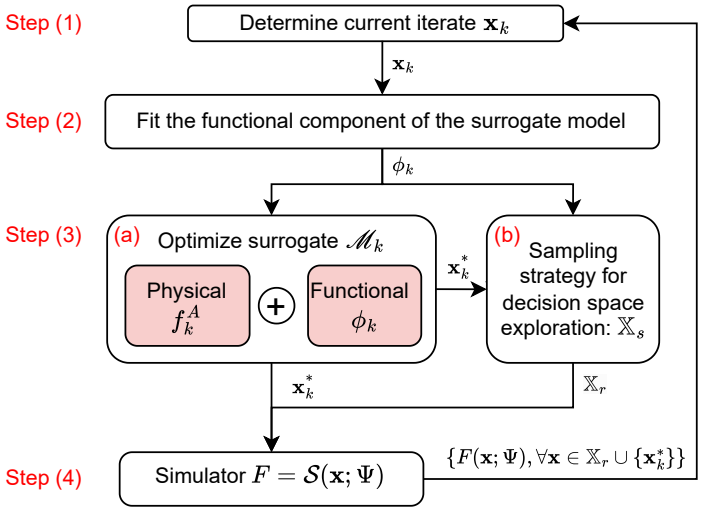


Figure 3.3: Surrogate model framework.

Figure 3.3 shows the main steps of the framework for building a reliable surrogate model. For a given iteration k, the framework will execute the following steps:

Step (1) The framework determines the best decision until iteration k as the *current iterate* ($\mathbf{x}_k$) according to their performance estimates resulting from simulations.

Step (2) Using all decision samples that have been simulated until iteration k to fit the functional component to estimate the parameters involved ($\boldsymbol{\beta}$). By doing this, the functional component is updated with the simulation samples added at iteration k.

Step (3) This step includes two processes, namely exploitation and exploration. (a) Exploitation: The surrogate model is updated by changing its functional component with the updated version resulting from Step (2). With the updated surrogate model, we solve the optimization problem under investigation to obtain the optimal (suboptimal) solution (solutions) under the current approximating state, namely $\mathbf{x}_k^*$. The initial guess often uses the current iterate $\mathbf{x}_k$. (b) Exploration: In order to prevent the algorithm from getting stuck in a local optimum, decision sampling is implemented to explore the decision space. Sometimes, this sampling step might be conducted only within a neighborhood of $\mathbf{x}_k^*$ to reduce unhelpful points.

Step (4) Apply traffic simulators to evaluate the decision samples obtained from Step (3) to augment the dataset of simulation samples. Terminate the procedure if the maximum number of simulation runs is reached, otherwise proceed to Step (1).

Essentially, the physical component is a macroscopic approximation of the traffic simulator. It is less detailed and accurate but is more computationally efficient to evaluate given its simplicity and differentiable property. The functional component, on the other hand, models the distance between the simulated data and the approximations derived from the physical component. It follows that the surrogate $\mathscr{M}$ can be interpreted as a macroscopic approximation to the objective function provided by f^A, which is corrected parametrically by a scaling factor η and a separable error term ϕ (Osorio and Chong, 2015). With algorithm iterating, the surrogate model $\mathscr{M}$ will become more and more accurate, i.e., the results from the surrogate model will approach the results from simulations. Therefore, it can be used to find the optimal solution under the current approximating state by solving an analytical optimization problem that is usually easy to solve with a variety of mainstream solvers. $\mathscr{M}$ thus can be used to approximate $\hat{f}$ and accelerate the algorithm convergence, thereby reducing the number of simulation runs needed to find well-performing solutions of the SO problem.

3.4 Data and study areas

This section summarizes the data used and study areas considered in the next chapters. Table 3.2 shows the chapters where these data and study areas are considered.

Table 3.2: Summary of data and study areas.

Method	Chapter
Data	
Foursquare location-based social network data	Chapter 4
Location-based social network data processing	Chapter 4
Google popular times data	Chapter 5
Google popular times data processing	Chapter 5
Study areas	
Tokyo city, Japan	Chapter 4
Kyoto city, Japan	Chapter 6, Chapter 7, Chapter 9
Munich city, Germany	Chapter 6, Chapter 7

3.4.1 Data

3.4.1.1 Foursquare location-based social network data

An LBSN check-in event is automatically recorded when a user posts with geo-location information or visits a venue (a point-of-interest). Each check-in is described by a user ID, a venue ID, and the time of the check-in. Table 3.3 lists the main fields of a check-in. In this regard, we can treat venues as detectors of check-in events, and users are the objects being detected. Compared to conventional household surveys, LBSN check-in data can be collected with a much higher frequency at a very low cost, and compared to traffic flow measurements, detectors of check-in events (i.e., venues) are "deployed" much denser within the urban area. The high resolution in both spatial and temporal dimensions has made LBSN data a reliable data alternative to model and estimate OD patterns within urban areas.

Table 3.3: Main fields of check-in events.

Field	Description	Example
Time	Time of the check-in	Apr 03, 2013, 18:17:18
User ID	Unique anonymized ID of the user	4bf58dd8d4898
Venue ID	Unique anonymized ID of the venue	4f0fd5a8e4b
Latitude	Latitude of the venue	35.7051
Longitude	Longitude of the venue	139.6196
Venue category	Category of the venue	Restaurant

We test the OD estimator presented in Chapter 4 using the Foursquare check-in data. Foursquare was launched in 2009 and has provided the leading LBSN service for more than a decade. As of 2019, it has included 105 million POIs in 190 countries and regions and received 1 billion check-ins annually. Foursquare data thus has broad spatial coverage and can somewhat capture human behavior in urban areas. Foursquare has its

own proprietary taxonomy of more than 1000 categories. According to the hierarchical taxonomy of categories (version 2012), ten parent categories are defined, including Arts & Entertainment, College & University, Event, Food, Nightlife Spot, Outdoors & Recreation, Professional & Other Places, Residence, Shops & Service, Travel & Transport.

3.4.1.2 Location-based social network processing

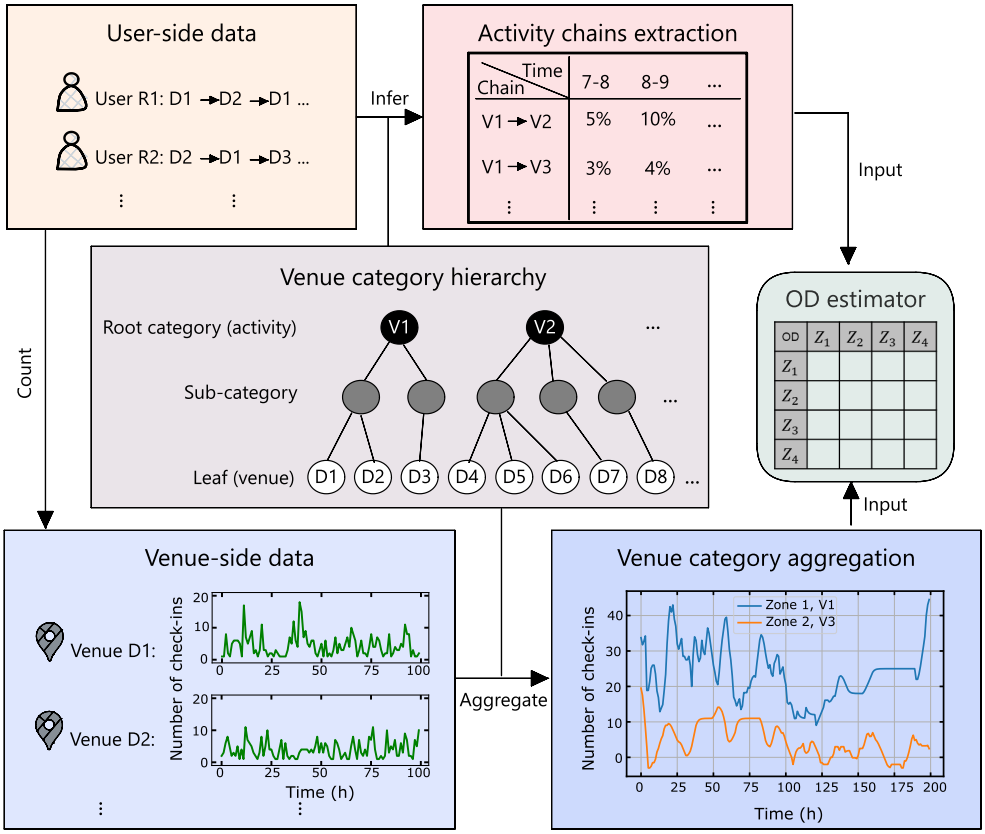


Figure 3.4: Process of extracting relevant information from LBSN check-in data for OD estimation.

By combining with the pre-registered location and category information of venues, check-in data has become a carrier of activity-oriented urban mobility patterns. Such data can be used to model and understand urban travel demand by aggregating the check-ins at venues based on specific categorization methods. Figure 3.4 depicts the data processing and information extraction procedure for preparing necessary inputs to the proposed OD estimator. Normally, venue-side data and user-side data are distinguished in the site server (Yang et al., 2015). Venue-side data contains the check-in statistics with respect to venues, while user-side data preserves the check-in history of users. As shown in the figure, venue-side data are obtained by aggregating the user-side data at the venue

level. It has been recognized that one can aggregate venue-side check-in data based on the categorical hierarchy adopted by the site to model the activity-based mobility flows (Hu et al., 2019). For example, in a three-level categorical hierarchy, Chinese restaurants and buffets belong to the venue type "Restaurant" which is a sub-category node of the root category "Food". Each root category can be viewed as one certain type of activity. The categorical hierarchy thus provides an intuitive way to group venues into activities. By comparing the patterns of different activities at different locations, one is able to analyze mobility patterns. On the other hand, the activity share matrix at different times of the day can be estimated by aggregating the activity-level movements deduced from the activity chains extracted from user-side data. While activity check-ins reflect the respective productions and attractions, activity share matrices can provide information for estimating the distributions for those productions and attractions. Therefore, we can combine the information contained in LBSN data to estimate and understand the mobility patterns within urban areas.

3.4.1.3 Google popular times data

Google's Popular time (GPT) graph (Google, 2023) displays the level of busyness of a Point of Interest (POI) at different times of the day relative to its busiest hour of the week. The historical busyness of a POI is measured on a relative scale of 0 to 100, with 100 indicating the busiest hour. The live busyness is computed based on the relative busyness compared to the corresponding historical period. The popular time graph for POIs is publicly accessible on Google Maps (Google Maps, 2023). Figure 3.5 provides an example of the Google Popular Times (GPT) graph for the Munich Central Station on April 7, 2023. The blue bars show the historical busyness levels hourly. The red bar represents the live busyness level when it is retrieved.

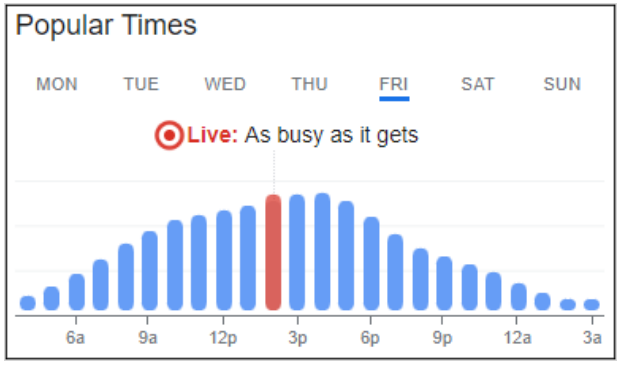


Figure 3.5: Example of GPT graph showing historical (blue) and live popularity (red).

The use of GPT graph as a crowdsensed data source for measuring the busyness of POIs, such as PT stations, has been recognized as a valuable approach for exploring urban dynamics (e.g., patterns of urban activity) (Vitello et al., 2023; Niu and Silva, 2020). Given the high penetration of Google services, this publicly accessible proxy for

crowding is available at a sheer scale in any city (Vitello et al., 2023; Möhring et al., 2021).

3.4.1.4 Google popular times data processing

GPT data is collected and used to measure demand pattern changes in PT stations in Germany in Chapter 5. Specifically, the GPT data of 2,134 railway stations, including U-Bahn (urban subway) and S-Bahn (city rapid railway) stations, were collected in two-hour intervals before (May 26 – May 31), during (Jun 1 – Aug 31) and after (Oct 9 – Dec 9) the 9-EUR implementation. Figure 3.6 shows the location distribution of these stations across Germany.

Considering that some stations may have very sparse entries, the collected GPT dataset is cleaned through three filters: (i) Historical busyness cannot be zero at all time intervals for an entire week; (ii) Historical busyness should be updated at least once a week; (iii) Live data of at least one time interval should be non-null in all periods before, during, and after the implementation of 9-EUR. Finally, 293 stations were preserved.

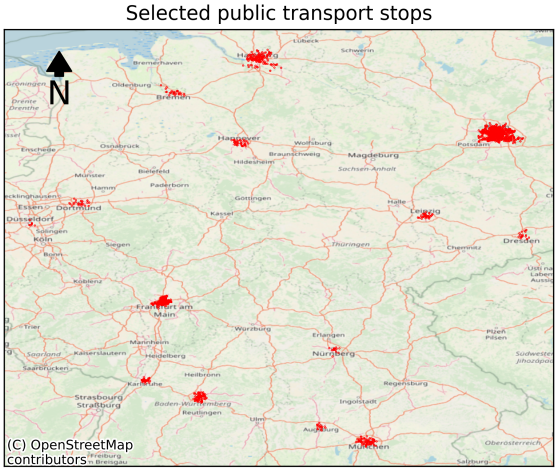


Figure 3.6: Location of selected PT stations.

3.4.2 Study areas

3.4.2.1 Tokyo city, Japan

The Foursquare check-in data of Tokyo city, Japan, from April 2012 to February 2013, are used in the experiments conducted in Chapter 4. We refer the reader to Yang et al. (2014a) for a more detailed description of the dataset. Figure 3.7a shows the map of the study area and the delineation of Traffic Analysis Zones (TAZs). The study area (1,302 km^2) is divided into 17 TAZs. Figure 3.7b exhibits a heatmap of 10,000 check-in records randomly sampled from the entire dataset which contains 573,703 records. The heatmap has a clear center and the color intensity gradually fades from the center outward.

(a) Layout of TAZs **(b)** Heatmap of check-ins

Figure 3.7: Study area: City of Tokyo.

Due to the lack of venue-side data, we aggregate user-side data hourly for each root venue category, each TAZ, and each day to reconstruct the venue-side dataset. Categories with fewer than five check-ins are not further defined as activity nodes of the TAZ. This can help identify the functionality of TAZs and the source of their production and attraction over time. It is worth mentioning that even though "home" and "work" activities are not covered adequately in LBSN data, the demand for them is still partially included in "Residential" and "Professional" activities, respectively.

Furthermore, we apply the moving average (seven days) technique to cancel some randomness of check-in behavior. In terms of the user-side dataset, we first extract the activity chain of each user. An activity share matrix can then be derived by counting the transfers between every two activities followed by normalization.

3.4.2.2 Kyoto city, Japan

The selected area covers the main city of Kyoto, enclosed by four main roads in different directions. The Kyoto study area spans approximately 6 km × 8 km (48 km^2). The network consists of 1189 links with 217 detectors. In this network, detectors are fairly evenly distributed, as shown in Figure 3.8.

At regular times of the year, various events are held in Kyoto, such as the Gion Festival[1] in July. During the Gion Festival, Shijo, the main road in the center of Kyoto, and the surrounding roads are temporarily controlled at night for about one to two weeks, during which motor vehicles are prohibited from entry. Figure 3.8c shows the area that might be closed temporarily during the festival. Temporary link/area closures due to such special events can render the supply-side disruptions (SSDs) discussed in Chapter 6. Therefore, the SSD scenario of the Kyoto network is defined as the closure of links highlighted in Figure 3.8c from 7:00 to 8:00, during which 85 links (7% of all) are

[1]See `https://www.yasaka-jinja.or.jp/event/gion/` for more information about the festival.

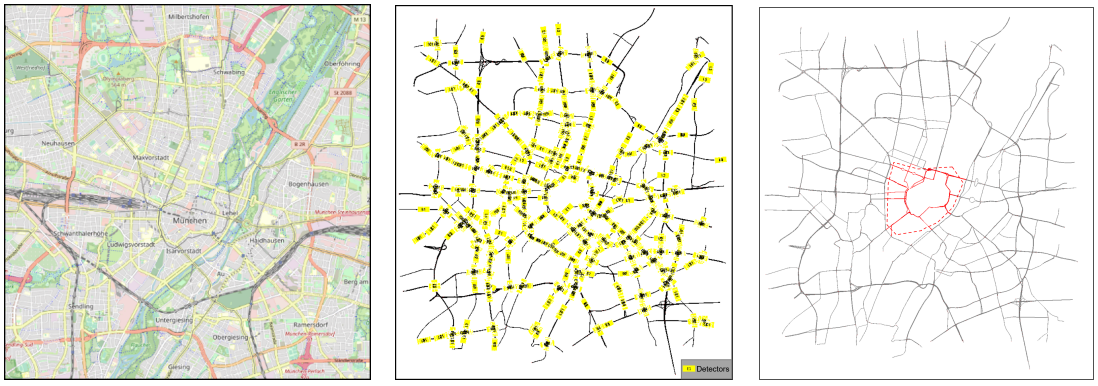
closed. All links are open outside this period. Due to a lack of required demand data, the closure is set in the morning rather than evening as in reality.

(a) Kyoto city center **(b)** Kyoto network detectors **(c)** Kyoto disruption area

Figure 3.8: Kyoto city center and locations of detectors.

3.4.2.3 Munich city, Germany

The concerning area of Munich is about 10 km × 10 km (100 km^2) large, covering the busiest streets and commercial areas in Munich. The network consists of 2605 links with 564 detectors. Detectors are evenly distributed across the network, as shown in Figure 3.9b. Figure 3.9c shows a supply-side disruption (SSD) scenario design for the comparative analysis conducted in Chapter 6 to examine the role of network topology in traffic resilience. The area closure indicated in Figure 3.9c serves as the hypothetical SSD scenario. Similarly, 209 links (8% of all) are closed from 7:00 to 8:00.



(a) Munich city center **(b)** Munich network detectors **(c)** Munich disruption area

Figure 3.9: Munich city center and locations of detectors.

Chapter 4

Origin-destination demand estimation with location-based social network data

The content of this chapter has been partially presented in the following work:

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2024b). A two-stage stochastic programming approach for dynamic OD estimation using LBSN data. *Transportation Research Part C: Emerging Technologies*, 158:104460.

4.1 Introduction and research contributions

Existing OD estimation methods primarily rely on three travel data sources, i.e., traditional household surveys, traffic measurements, and positioning technology-based data (Yang et al., 2015). Traditional household surveys are time-consuming, labor-intensive, and expensive, therefore are normally restrained within a limited area at a low frequency (e.g., once or twice a decade). These surveys, while providing detailed socio-demographical representation of demand, are only viable to develop planning models that depict average network conditions due to their limited frequency. The second data type of traffic measurement relies on fixed detection infrastructure distributed over the network and has been widely used in DODE methods since it provides the required time-varying dynamics of network states for estimation. However, traffic-measurement-based DODE methods, on one hand, structurally suffer from the issue of indeterminateness in estimating realistic OD flows patterns (i.e., multiple sets of varying OD matrix patterns can satisfy the constraints imposed by the traffic measurements and optimize the system objective at the same time) (Cascetta et al., 2013; Antoniou et al., 2016; Qurashi et al., 2022) and on the other require significant computational resources to run dynamic traffic simulations that map estimated demand patterns on network models to attain and match traffic measurements iteratively.

Methods using the third data source type have attracted much attention in recent years. The ubiquity of smartphones equipped with positioning technologies, such as GPS and Bluetooth, has resulted in the regular real-time generation of large sets of well-distributed data that also provides unprecedented opportunities for the implementation and application of OD estimation methods. One such suitable data type is LBSN data that has been used to develop demand models, due to its broad urban spatial and temporal coverage and confirmed trip purposes (Hu and Jin, 2017). LBSN services

generate a large amount of anonymous check-in data of venues and users, making it a natural "host" of urban mobility patterns (Jin et al., 2014; Yang et al., 2015). More specifically, the check-in time series of venues records the travel destination distribution in both spatial and temporal dimensions, while the check-in history of users reflects the activity chains of individuals. Utilizing the LBSN data allows replacing the traffic measurements with the check-in data, which on one hand is a more direct and accurate representation of the demand patterns and on the other removes the significant burden of developing traffic models and estimating simulation-based objective functions.

To tap the potential of using LBSN data for dynamic demand estimation, this chapter establishes a TSSP framework integrating the activity chains to model activity-level mobility flows using LBSN data. To the best of our knowledge, this is the first effort to apply it to model the dynamic OD estimation problem based on LBSN data.

The contributions of this chapter are as follows:

- We develop a dynamic OD demand estimation method based on the TSSP framework using a type of emerging opportunistic data, location-based social network data, which supplements the body of OD estimation algorithms, especially under transportation disruptions. This approach can effectively mitigate the indeterminateness issue inherent in methods based on traffic measurements.
- We develop a GBD algorithm to efficiently solve the constructed OD estimation problem, enabling the application of the proposed estimator to online estimation.

This chapter is structured as follows: In Section 4.2, the mathematical model of the OD estimator based on LBSN data is constructed, followed by the solution algorithms. Later on, case studies are elaborated and model performance is evaluated. Finally, we draw some conclusions and suggest future directions for research.

4.2 Methodology

4.2.1 Traffic analysis zone graph model

To facilitate the subsequent model development, we introduce the graph model of a TAZ as depicted in Figure 4.1. Specifically, the nodes within the graph model include a group of activity nodes, a virtual source, and a virtual sink. The edges, on the other hand, comprise activity flows that connect the activity nodes and virtual flows that link the virtual source and sink to the activity nodes. Note that inter-zone flows, connecting different TAZs, fall outside the scope of any specific TAZ. Activity nodes and activity flows are defined as follows:

Definition 1 (Activity node) *An activity node is a concentrating representation of the venues belonging to a certain root category inside the modeled traffic analysis zone.*

Definition 2 (Activity flow) *An activity flow indicates the movements of people between two activity nodes within the modeled traffic analysis zone.*

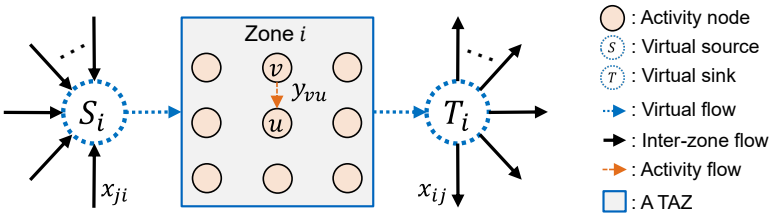


Figure 4.1: Graph model of a traffic analysis zone.

For example, of the venue category hierarchy illustrated in Figure 3.4, all venues falling under the category V_1 located within a given TAZ collectively form an activity node. The activity flows represent the movements between venue categories. In addition, we create a virtual source and a virtual sink for each TAZ for respectively aggregating the total incoming and outgoing trips, and effectively bridging the inter-zone flows and zonal activity flows. Here, inter-zone flows are what we traditionally call OD flows. In the following, we will also refer to them interchangeably as either inter-zone flows or OD flows. Virtual sources and virtual sinks are defined as follows:

Definition 3 (Virtual source) *A virtual source of a traffic analysis zone refers to a node that serves as the central recipient for all inter-zone flows emanating from other traffic analysis zones. The virtual source is responsible for allocating these inflows across all activity nodes within the same zone.*

Definition 4 (Virtual sink) *A virtual sink of a traffic analysis zone refers to a node that serves as the central origin for all inter-zone flows destined for other traffic analysis zones. The virtual sink is responsible for allocating the outflows from the zonal activity nodes to other zones.*

For a TAZ graph model, its virtual source and sink are connected to all of its activity nodes. Virtual flows are used to represent the movement of individuals either from the virtual source to an activity node or from an activity node to the virtual sink. Thus, the sum of flows to the sink and the sum of flows to the source can represent the zonal trip production and attraction, respectively. Later, we will show that inter-zone flows and activity flows are the first-stage and second-stage decision variables in the proposed TSSP model. That is to say, virtual sources and sinks are the bridge that links the first-stage and second-stage decisions. This treatment also enables the utilization of decomposition algorithms to expedite the solution of the OD estimation problem.

4.2.2 LBSN-data-based OD estimator

OD flows are generated when people travel for various activities across different locations. Namely, OD patterns are the aggregated result of activity check-in patterns. Since OD patterns of a certain time-of-day interval would not change significantly within a period free of disruptive events (i.e., without critical changes in either supply or demand), the

following assumption is plausible and is introduced to drive the modeling of the LBSN-data-based OD estimator.

Assumption 1 *For a certain time-of-day interval, similar activity check-in patterns on different days within a predefined reference period are generated by similar OD patterns.*

With this assumption, a certain OD matrix needs to be mapped onto multiple check-in pattern scenarios (defined later in this section). In essence, this assumption allows us to exert additional constraints on the OD estimation problem for a specific time-of-day interval, thereby reducing the search space of the posterior OD estimates. Therefore, the LBSN-data-based OD estimator holds the potential of mitigating the indeterminateness issue existing in those based on traffic measurements.

The previous analysis inspires the development of an LBSN-data-based dynamic OD estimator leveraging a two-stage stochastic programming framework for incorporating multiple check-in pattern scenarios. Stochastic programming is an approach for modeling optimization problems under uncertainty. In the proposed model, check-in patterns represent the uncertain output of a certain OD matrix input to the system. In the two-stage framework, OD flows are the decision variables of the first-stage problem, and activity flows are the decisions of the second-stage problem.

To formulate the problem, we present the following modeling process for a specific time interval τ. For ease of presentation, the superscript τ will be omitted unless specified. With a slight abuse of notation, the variables corresponding to $\tau-1$ will be indicated with a superscript (l). For example, $\mathbf{q}^{\tau}$ and $\mathbf{q}^{\tau-1}$ will be written as $\mathbf{q}$ and $\mathbf{q}^{(l)}$, respectively. Some important notations with their descriptions are listed in Table 4.1.

4.2.2.1 First-stage problem: OD flows estimation

In the first stage, the OD flows are the decision variables. They are determined within the space restricted by three key considerations. First, akin to traffic-measurement-based OD estimators, taking into account the deviation between the posterior OD estimate and the prior OD estimate in the objective function can align the posterior OD estimate with a known OD pattern. The prior OD flows can be historical values or results from a four-step model that can capture the OD patterns under similar conditions. Second, we can easily deduce the zonal trip production and attraction for every TAZ from the estimated OD flows. Combining these estimated production and attraction values with a relationship model between zonal check-in counts and zonal production/attraction is capable of adjusting the OD demand level based on check-in observations. Third, the inclusion of multiple check-in pattern scenarios in the framework also plays a critical role in constraining the search space of the posterior OD flows.

Therefore, we formulate the generic objective function of the two-stage stochastic programming model for OD estimation using LBSN check-in data as follows:

$$\min_{\mathbf{x}} \quad w_x f_x\big(\mathbf{x}, \mathbf{x}^{(p)}\big) + w_p f_p\big(\hat{P}(\mathbf{x}), P(\mathbf{q})\big) + w_a f_a\big(\hat{A}(\mathbf{x}), A(\mathbf{q})\big) + w_s \mathbb{E}_{\xi}[Q(\mathbf{x}, \xi)] \tag{4.1}$$

Table 4.1: Notation in the LBSN-based OD estimator.

Sets	
$\mathbb{V}_i$	set of activity nodes in the graph model of TAZ i
$\mathbb{Z}$	set of TAZs in the study area
Parameters	
$\boldsymbol{\rho}$	activity share matrix
$\rho_{uv,i}$	activity share from activity node u to v in i (normalized by the check-ins at u)
$\mathbf{x}^{(p)}$	prior OD estimates
$\mathbf{q}$	observed check-in counts
$q_{v,i}$	observed check-in counts at activity v within TAZ i
$\mathbf{q}^{(l)}$	observed check-in counts in the last time interval of the interval of interest
$q_{v,i}^{(l)}$	observed check-in counts at activity v within TAZ i in the previous time interval
$\hat{P}(\mathbf{x})$	zonal production derived from $\mathbf{x}$
$\hat{A}(\mathbf{x})$	zonal attraction derived from $\mathbf{x}$
$P(\mathbf{q})$	zonal production estimated based on the observed check-in counts $\mathbf{q}$
$A(\mathbf{q})$	zonal attraction estimated based on the observed check-in counts $\mathbf{q}$
ξ	problem state at the second stage representing the second-stage scenario
$q_{v,i}(\xi)$	check-in counts of activity v in TAZ i under scenario ξ
$\Delta_i(\xi)$	difference of check-in counts in scenario ξ between two successive time intervals
$\hat{\Delta}_i(\mathbf{y}_i)$	difference of check-in counts derived from the optimized activity flows $\mathbf{y}_i$
Y	observed traffic measurements
$\mathcal{A}$	traffic assignment method
w_x	weighting factor for prior OD estimates in the objective function
w_p	weighting factor for zonal productions in the objective function
w_a	weighting factor for zonal attractions in the objective function
$\underline{\epsilon}_b$	threshold parameter of the lower bound for the posterior OD flows
$\overline{\epsilon}_b$	threshold parameter of the upper bound for the posterior OD flows
$\underline{\epsilon}_a$	threshold parameter of the lower bound for activity shares
$\overline{\epsilon}_a$	threshold parameter of the upper bound for activity shares
$\underline{\epsilon}_s$	threshold parameter of the lower bound for the source imbalance
$\overline{\epsilon}_s$	threshold parameter of the upper bound for the source imbalance
$\underline{\epsilon}_t$	threshold parameter of the lower bound for the sink imbalance
$\overline{\epsilon}_t$	threshold parameter of the upper bound for the sink imbalance
N_s	number of second-stage scenario samples
Decision variables	
x_{ij}	OD (inter-zone) flow from TAZ i to j
$y_{vu,i}$	activity flow from activity node v to u within TAZ i

where $\mathbf{x}$ is the decision variable of the first-stage problem, i.e., OD flows, $\mathbf{x}^{(p)}$ is the given prior OD flows. $\mathbf{q}$ is the vector of observed check-in counts, with each element representing the check-in counts at an activity node. $f_x(\cdot)$ is the Goodness-of-Fit (GoF) function measuring the difference between the posterior and prior OD flows. $\hat{P}(\mathbf{x})$ and $\hat{A}(\mathbf{x})$ are the vectors of out-flows (i.e., zonal production) and in-flows (i.e., zonal attraction) of zones, which are obtained by aggregating the posterior OD flows $\mathbf{x}$ correspondingly. $P(\mathbf{q})$ and $A(\mathbf{q})$ are the given production and attraction vectors estimated with the observed check-in counts $\mathbf{q}$. The empirical findings about their relationship are expounded on in Section 4.3.1. $f_p(\cdot)$ and $f_a(\cdot)$ measure the GoF between the modeled and the measured zonal production and attraction, respectively. w_x, w_p and w_a are weighting factors that quantify the relative reliability of the prior OD estimate, the prior production estimate and the prior attraction estimate. The last term in Equation (4.1) measures the expected minimum difference between the estimated and observed check-in counts difference given the OD flows $\mathbf{x}$. w_s is a weighting factor that quantifies the trade-off between the optimization of OD flow patterns and check-in patterns.

Besides, we enforce bound constraints on the OD flows to prevent the emergence of unrealistic solutions. These bounds are defined as multiples of the prior OD estimates given by

$$\underline{\epsilon}_b x_{ij}^{(p)} \leq x_{ij} \leq \bar{\epsilon}_b x_{ij}^{(p)} \qquad \forall i, j \in \mathbb{Z} \tag{4.2}$$

Where x_{ij} denotes the OD flow from TAZ i to j, $\mathbb{Z}$ denotes the set of TAZs under consideration, $\underline{\epsilon}_b$ (< 1) and $\bar{\epsilon}_b (> 1)$ are threshold parameters.

4.2.2.2 Second-stage problem: Activity flows estimation

The further restriction of search space of OD flows is achieved by introducing a set of check-in scenarios of the second-stage problem state, as expressed by the last term in Equation (4.1), i.e., $w_s \mathbb{E}_\xi[Q(\mathbf{x}, \xi)]$. $\mathbb{E}_\xi$ calculates the expectation with respect to a random vector ξ, defined on the probability space $(\Omega, \mathscr{F}, \mathcal{P})$, with Ω being the sample space, $\mathscr{F}$ being the event space, and $\mathcal{P}$ being a probability distribution defined on $\mathscr{F}$. ξ is a random variable describing the problem state at the second stage. $Q(\mathbf{x}, \xi)$ is the optimal value of the second-stage problem under scenario ξ.

In this study, a realization of the second-stage problem represents a check-in pattern scenario. Here, check-in patterns represent the changes in check-in counts between two successive time intervals. Considering the activity share difference in different times-of-day intervals, a check-in pattern scenario can then be described as a tuple composed of the check-in counts of the relevant two time intervals and the corresponding activity share matrix, i.e., $\left(\mathbf{q}^{(l)}, \mathbf{q}, \boldsymbol{\rho}\right)$, where $\boldsymbol{\rho}$ indicates the activity share matrix at τ.

The objective of the second-stage problem is to determine the activity flows that minimize the deviation between the estimated check-in counts difference and the observed one (i.e., $\mathbf{q}^{(l)} - \mathbf{q}$). Noting the structure of the proposed TAZ graph model and the definition of check-in pattern scenarios, three conditions must be taken into account and modeled in this stage, including (i) the conservation of check-in counts at activity nodes, (ii) the shares of activity flows, and (iii) the balance of inter-zone OD flows and

zonal activity flows at the virtual sources and sinks. For illustration purposes, Figure 4.2 provides a graphical presentation of these conditions.

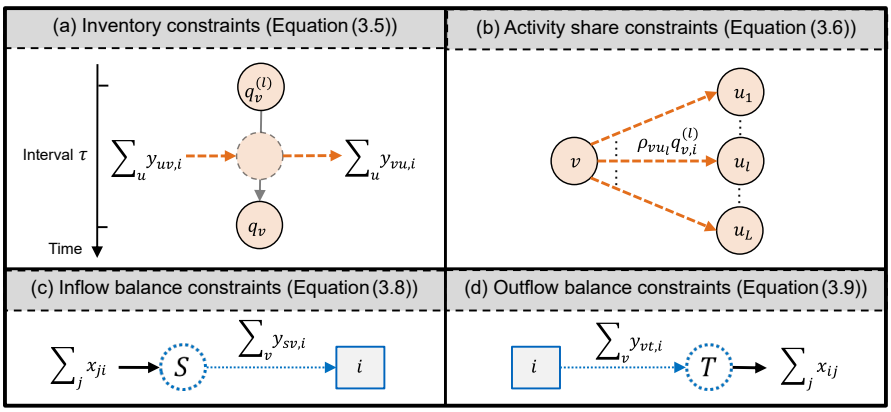


Figure 4.2: Graphic illustration of the conditions considered in the second-stage problem.

For a given check-in scenario ξ, the objective function of the second-stage problem can be expressed as

$$Q(\mathbf{x}, \xi) = \min_{\mathbf{y}} \quad \sum_{i \in \mathbb{Z}} f_s\big(\hat{\Delta}_i(\mathbf{y}_i), \Delta_i(\xi)\big) \tag{4.3}$$

where $\Delta_i(\xi) = \mathbf{q}_i^{(l)}(\xi) - \mathbf{q}_i(\xi)$ is the observed check-in counts difference in scenario ξ, $\hat{\Delta}_i(\mathbf{y}_i)$ is the estimated check-in counts difference derived from the optimized activity flows $\mathbf{y}_i$. As can be seen, this objective function is also the term inside the expectation calculator in the last term of the first-stage objective.

For activity node v in TAZ i, we have $\hat{\Delta}_{v,i} = \sum_{u \in (\mathbb{V}_i \setminus \{v\}) \cup \{t\}} y_{vu,i} - \sum_{u \in (\mathbb{V}_i \setminus \{v\}) \cup \{s\}} y_{uv,i}$. $f_s(\cdot)$ is a GoF function measuring the fitting between the estimated and observed check-in counts differences.

The first condition (see Figure 4.2a) expresses that for an activity node v, the total departing flows cannot be greater than the combined arriving flows and the number of check-ins recorded during the previous interval. It can be written as

$$\sum_{u \in (\mathbb{V}_i \setminus \{v\}) \cup \{t\}} y_{vu,i} - \Big(\sum_{u \in (\mathbb{V}_i \setminus \{v\}) \cup \{s\}} y_{uv,i} + q_{v,i}^{(l)}(\xi) \Big) \leq 0 \qquad \forall v \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.4}$$

where $\mathbb{V}_i$ is the set of activity nodes in TAZ i. In practice, only the main venue categories in the TAZ will be considered for the sake of: (i) reducing the noise in the statistics caused by insufficient venues of a specific category; (ii) distinguishing different TAZs with respect to the land-use functionality and characteristics.

We denote $\rho_{vu,i}$ the share of activity flow from activity node v to u within TAZ i during interval τ. We can thus construct activity share constraints for the second condition (see Figure 4.2b) to restrict the search space of activity flows by making use of the activity

share matrix. This matrix encapsulates the activity chain information derived from user-side data. The constraints are given by

$$(1-\underline{\epsilon}_a)\rho_{vu,i}q_{v,i}^{(l)} \leq y_{vu,i} \leq (1+\overline{\epsilon}_a)\rho_{vu,i}q_{v,i}^{(l)} \qquad \forall v,u \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.5}$$

where $\underline{\epsilon}_a$ and $\overline{\epsilon}_a$ are predefined threshold parameters within the range (0,1). They serve the purpose of mitigating the over-fitting issue during optimization by permitting slight deviations of activity flows from their theoretical values. Note, the activity share matrix can be normalized at either the network level or the TAZ level. For simplification, for a specific time-of-day interval of the day, we consider a common activity share matrix for the entire area, which can be estimated from the historical check-in data. As such, $\rho_{vu,i} = \rho_{vu}, \forall i \in \mathbb{Z}$. Then, Equation 4.5 can be simplified as

$$(1-\underline{\epsilon}_a)\rho_{vu}q_{v,i}^{(l)} \leq y_{vu,i} \leq (1+\overline{\epsilon}_a)\rho_{vu}q_{v,i}^{(l)} \qquad \forall v,u \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.6}$$

Regarding the third condition, it is natural to impose the inflow and outflow balance constraints on the source and sink nodes, as shown in Figure 4.2c and 4.2d, respectively. Importantly, this condition establishes a connection between the decision variables in the first- and second-stage problems. More specifically, the inflow balance indicates that, for a specific TAZ, the aggregate virtual flows from the source node should equal the sum of inter-zone flows to it. Similarly, the outflow balance indicates that, for a specific TAZ, the sum of virtual flows to the sink node should match the sum of inter-zone flows from it. Nevertheless, due to the randomness and incompleteness of the activity information, it is reasonable allow a subtle deviation in both constraints. Consequently, the inflow and outflow balance can be expressed as

$$(1-\underline{\epsilon}_s)\sum_{j\in\{\mathbb{Z}-i\}} x_{ji} \leq \sum_{v\in\mathbb{V}_i} y_{sv,i} \leq (1+\overline{\epsilon}_s)\sum_{j\in\{\mathbb{Z}-i\}} x_{ji} \qquad \forall i \in \mathbb{Z} \tag{4.7}$$

$$(1-\underline{\epsilon}_t)\sum_{j\in\{\mathbb{Z}-i\}} x_{ij} \leq \sum_{v\in\mathbb{V}_i} y_{vt,i} \leq (1+\overline{\epsilon}_t)\sum_{j\in\{\mathbb{Z}-i\}} x_{ij} \qquad \forall i \in \mathbb{Z} \tag{4.8}$$

where $\underline{\epsilon}_s$, $\overline{\epsilon}_s$, $\underline{\epsilon}_t$ and $\overline{\epsilon}_t$ are predefined threshold parameters in the range (0,1), representing the extent to which the conservation can be violated. $y_{sv,i}$ denotes the virtual flow from the virtual source to activity node v within TAZ i, while $y_{vt,i}$ denotes the virtual flow from v to the virtual sink.

Furthermore, all activity flows and virtual flows should be non-negative as expressed by

$$y_{vu,i} \geq 0, y_{sv,i} \geq 0, y_{vt,i} \geq 0 \qquad \forall v,u \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.9}$$

Note that in the second-stage problem, the optimal activity flows $\mathbf{y}^*$ depend on the first-stage OD flows $\mathbf{x}$ and the second-stage problem state ξ.

4.2.2.3 Full two-stage stochastic model

By integrating the objective functions and constraints of the two stages of problems described above, we obtain the two-stage stochastic programming model of the LBSN-data-based OD estimation problem under uncertainty with multiple check-in pattern scenarios as below.

$$\min_{\mathbf{x}} \quad w_x f_x\big(\mathbf{x}, \mathbf{x}^{(p)}\big) + w_p f_p\big(\hat{P}(\mathbf{x}), P(\mathbf{q})\big) + w_a f_a\big(\hat{A}(\mathbf{x}), A(\mathbf{q})\big) + w_s \mathbb{E}_\xi[Q(\mathbf{x}, \xi)] \tag{4.10}$$

$$\text{s.t.} \quad \underline{\epsilon}_b x_{ij}^{(p)} \leq x_{ij} \leq \overline{\epsilon}_b x_{ij}^{(p)} \qquad \forall i, j \in \mathbb{Z} \tag{4.11}$$

where:

$$Q(\mathbf{x}, \xi) = \min_{\mathbf{y}} \quad \sum_{i \in \mathbb{Z}} f_s\big(\hat{\Delta}_i(\mathbf{y}_i), \Delta_i(\xi)\big) \tag{4.12}$$

$$\text{s.t.} \quad \sum_{u \in (\mathbb{V}_i \setminus \{v\}) \cup \{t\}} y_{vu,i} - \sum_{u \in (\mathbb{V}_i \setminus \{v\}) \cup \{s\}} y_{uv,i} - q_{v,i}^{(l)}(\xi) \leq 0 \quad \forall v \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.13}$$

$$(1 - \underline{\epsilon}_a)\rho_{vu} q_{v,i}^{(l)}(\xi) \leq y_{vu,i} \leq (1 + \overline{\epsilon}_a)\rho_{vu} q_{v,i}^{(l)}(\xi) \qquad \forall v, u \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.14}$$

$$(1 - \underline{\epsilon}_s) \sum_{j \in \{\mathbb{Z} - i\}} x_{ji} \leq \sum_{v \in \mathbb{V}_i} y_{sv,i} \leq (1 + \overline{\epsilon}_s) \sum_{j \in \{\mathbb{Z} - i\}} x_{ji} \quad \forall i \in \mathbb{Z} \tag{4.15}$$

$$(1 - \underline{\epsilon}_t) \sum_{j \in \{\mathbb{Z} - i\}} x_{ij} \leq \sum_{v \in \mathbb{V}_i} y_{vt,i} \leq (1 + \overline{\epsilon}_t) \sum_{j \in \{\mathbb{Z} - i\}} x_{ij} \quad \forall i \in \mathbb{Z} \tag{4.16}$$

$$y_{vu,i} \geq 0 \qquad \forall v, u \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.17}$$

$$y_{sv,i} \geq 0 \qquad \forall v \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.18}$$

$$y_{vt,i} \geq 0 \qquad \forall v \in \mathbb{V}_i, \forall i \in \mathbb{Z} \tag{4.19}$$

By comparison, the proposed LBSN-data-based OD estimator resembles the generic traffic-measurement-based OD estimator to a certain extent: (i) the proposed model also relies on a prior OD flow estimate; (ii) $f_p(\cdot)$ and $f_a(\cdot)$ in the objective function of the first-stage problem (Equation (4.10)) and the activity share constraints in Equation (4.14) play a similar role to a traffic assignment method; (iii) both models need to handle the stochasticity of the observed data.

On the other hand, the difference between the two methods lies in that LBSN data are collected when the travel is finished, but traffic measurements are collected during the travel. In other words, LBSN contains end-to-end information, while traffic measurements record the situation between ends. As a result, the application of LBSN-data-based OD estimators usually demands no network structure but needs information on the activity preferences of travelers. More importantly, different from most traffic-measurement-based OD estimators in the literature, the proposed model can be used for dynamic OD estimation by using only the LBSN data without the need to run computationally expensive dynamic traffic simulations to generate simulated traffic measurements. This puts the proposed methodology at a significant computational advantage against most dynamic OD estimation approaches. Moreover, unlike the previous LBSN-based estimators, essentially, our model derives the OD matrix by reconstructing the activity-based mobility flows, making use of the confirmed trip purpose information.

4.2.3 Solution algorithms

To solve this model, the SAA and MCMC sampling techniques introduced in Section 3.3.1.1 and Section 3.3.2.1, respectively, are integrated to generate reliable check-in pattern scenarios. Then, the GBD algorithm provided in Section 3.3.1.2 can be applied to estimate the OD demand. It is worth mentioning that sampling techniques have to be fed by the check-in pattern scenario of the time interval under estimation (i.e., $(\mathbf{q}^{(l)}, \mathbf{q}, \boldsymbol{\rho})$) in order to generate reliable sampled scenarios. We call the scenario $(\mathbf{q}^{(l)}, \mathbf{q}, \boldsymbol{\rho})$ the reference scenario to distinguish it from the sampled scenarios. The sampled scenarios can be viewed as for the same time-of-day interval over different historical days that share similar context with the reference scenario.

We present an outline of the procedure employed to address the TSSP problem for LBSN-data-based OD estimation in Figure 4.3 to facilitate understanding. Given the required model inputs (i.e., $\mathbf{q}$, $\mathbf{q}^{(l)}$, $\boldsymbol{\rho}$, and $\mathbf{x}^{(p)}$) pertaining to the specific time interval of interest τ, MCMC sampling is utilized to generate a batch of similar check-in pattern scenarios based on the check-in patterns provided (i.e., $\mathbf{q}^{(l)}$ and $\mathbf{q}$). By integrating with the given activity shares information (i.e., $\boldsymbol{\rho}$), each individual scenario can be used to configure a corresponding second-stage problem. Combining these second-stage problems with the first-stage problem will result in the TSSP problem, which aims to estimate the OD flows for the given time interval. Subsequently, we employ the GBD algorithm to decompose the problem into a primary master problem and a series of subproblems. By solving these decomposed problems alternately, the optimal OD flows (i.e., $\mathbf{x}$) and activity flows (i.e., $\mathbf{y}$) that correspond to the observed check-in patterns will be obtained.

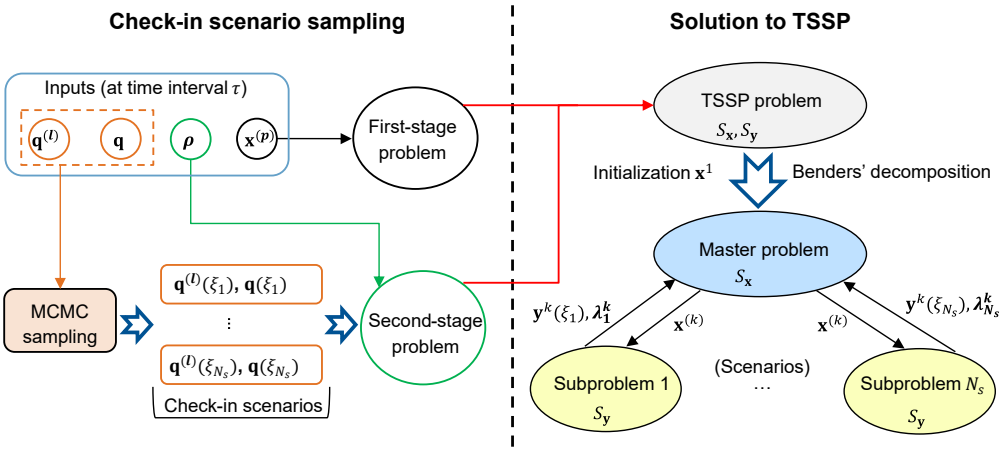


Figure 4.3: Solution procedure of the proposed LBSN-data-based OD estimator.

4.3 Case study and results

We conduct a case study in Tokyo city, Japan, by using Foursquare check-in data. The description and processing procedure of the Foursquare check-in data are provided

in Section 3.4.1.1 and Section 3.4.1.2, respectively. The delineation of TAZs and the distribution of check-in records across Tokyo are given in Section 3.4.2.1.

4.3.1 Algorithm setup

In the following experiments, we set the threshold parameters, $\{\underline{\epsilon}_s, \overline{\epsilon}_s, \underline{\epsilon}_t, \overline{\epsilon}_t, \underline{\epsilon}_a, \overline{\epsilon}_a\}$, as 0.2. The bound constraint parameters, $\{\underline{\epsilon}_b, \overline{\epsilon}_b\}$, are 0.2 and 5, respectively. We set the weighting factors $\{w_x, w_p, w_a, w_s\}$ as $\{1, 10, 0, 1\}$ unless otherwise specified.

For the MCMC sampling, the number of second-stage realizations N_s is set to 5. We assume a Gaussian distribution for the proposal distribution of an activity node v at time τ denoted by $\mathcal{S}_v(\cdot|q_v(\xi_n))$. Specifically, $\mathcal{S}_v(\cdot|q_v(\xi_n)) = \mathcal{N}(q_v(\xi_n), 0.1q_v(\xi_n))$, where $\mathcal{N}(\mu, \sigma)$ represents a Gaussian distribution with μ mean and σ^2 variance. Since Gaussian distributions are symmetric, $\mathcal{S}_v(q_v(\xi_n)|q_v(\zeta_n)) = \mathcal{S}_v(q_v(\zeta_n)|q_v(\xi_n))$, therefore Equation (3.16) reduces to

$$a_v(q_v(\xi_n), q_v(\zeta_n)) = \min\left\{1,\ \frac{\pi_v(q_v(\zeta_n))}{\pi_v(q_v(\xi_n))}\right\} \tag{4.20}$$

Regarding the GBD algorithm, the convergence threshold is $\epsilon = 0.05$. Without loss of generality, we apply the GLS estimator in the goodness-of-fit functions, $f_x(\cdot), f_p(\cdot), f_a(\cdot)$ and $f_s(\cdot)$, resulting in a convex optimization problem with complicating variables. In this case, the GBD algorithm can guarantee the same global optimum solution as the original problem. Mathematically, $f_x(\cdot), f_p(\cdot)$, $f_a(\cdot)$ and $f_s(\cdot)$ are given as follows:

$$f_x\big(\mathbf{x}, \tilde{\mathbf{x}}^\tau\big) = \big(\mathbf{x} - \tilde{\mathbf{x}}^\tau\big)^\top \Lambda_x \big(\mathbf{x} - \tilde{\mathbf{x}}^\tau\big) \tag{4.21}$$

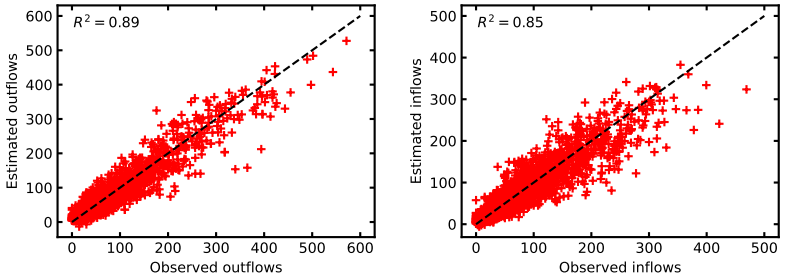
$$f_p\big(\hat{P}(\mathbf{x}), P(\mathbf{q})\big) = \big(\hat{P}(\mathbf{x}) - P(\mathbf{q})\big)^\top \Lambda_p \big(\hat{P}(\mathbf{x}) - P(\mathbf{q})\big) \tag{4.22}$$

$$f_a\big(\hat{A}(\mathbf{x}), A(\mathbf{q})\big) = \big(\hat{A}(\mathbf{x}) - A(\mathbf{q})\big)^\top \Lambda_a \big(\hat{A}(\mathbf{x}) - A(\mathbf{q})\big) \tag{4.23}$$

$$f_s\big(\hat{\Delta}_i(\mathbf{y}_i), \Delta_i(\xi_n)\big) = \big(\hat{\Delta}_i(\mathbf{y}_i) - \Delta_i(\xi_n)\big)^\top \Lambda_s \big(\hat{\Delta}_i(\mathbf{y}_i) - \Delta_i(\xi_n)\big) \tag{4.24}$$

where $\Lambda_x, \Lambda_p, \Lambda_a$ and Λ_s are the dispersion matrices of the prior OD estimates, out-flows distribution, in-flows distribution, and check-in pattern, respectively. For simplicity, we set $\Lambda_x = \Lambda_p = \Lambda_a = \Lambda_s = \mathrm{diag}(\mathbf{1})$.

Moreover, considering the linear relationship between the out-flows and the number of check-ins within the same TAZ observed from empirical data (as shown in Figure 4.4a), we calculate $P(\mathbf{q}) = \hat{\theta}^\top \mathbf{q}$, where $\hat{\theta} = (C^\top C)^{-1} C^\top P_0$, C is the matrix of the number of check-ins aggregated by TAZs, and P_0 is the vector of historical observed outflow patterns, as the reference to penalize the demand level deduced from the estimated OD flows. Figure 4.4a compares the observed out-flows and the out-flows estimated by a simple linear regression model based on the number of check-ins. The R-square is about 0.89. It is worth mentioning that the entire check-in dataset is used for estimating $\hat{\theta}$. Similarly, we can get the relationship model between in-flows and the number of check-ins $A(\mathbf{q})$ with the same method.



(a) Out-flows comparison **(b)** In-flows comparison

Figure 4.4: Linear relationship between out-flows (left), in-flows (right), and the number of check-ins.

4.3.2 Demand scenario setup

We conduct experiments on the morning peak (7 am - 10 am) of February 1st, 2013. The morning peak is divided into three estimation time intervals each for one hour. A one-hour time interval is set due to the limitation on the check-in data density of this dataset. However, the proposed methodology is applicable to smaller time intervals in the context of sufficient data, such as 15 minutes or half an hour. Regarding demand scenario generation, Antoniou et al. (2016) points out that the quality of the prior OD estimate, in terms of both demand level and patterns, is a key element that affects the performance of the OD estimator. Following the suggestions therein, we perturb the true OD flows to derive the historical OD flow estimates. More specifically, we create the prior OD estimate using the following formulation:

$$\tilde{\mathbf{x}}^{\tau} = (\phi_{rd} + \phi_{rm}\delta)\mathbf{x} \tag{4.25}$$

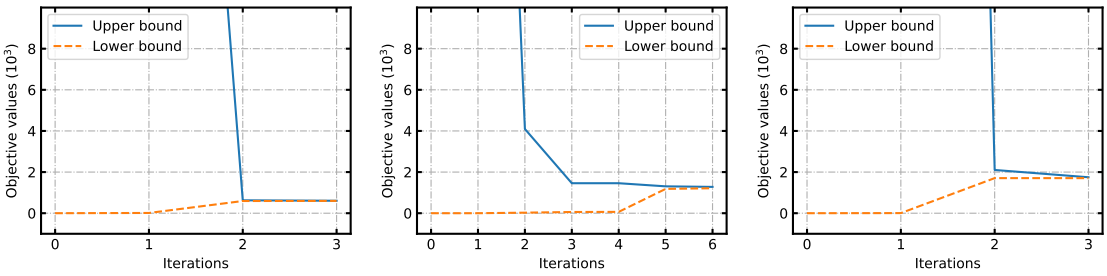
where ϕ_{rd} and ϕ_{rm} are the reduction and randomization parameters for perturbation, respectively, and δ is the random perturbation vector following Gaussian distribution. In the following experiments, we apply $\phi_{rd} = 0.7$, $\phi_{rm} = 0.3$, and $\delta \sim \mathcal{N}(0, 1/3)$ (99.7% of values located in [-1,1]), implying that the prior OD estimate is an out-of-date low-demand scenario.

4.3.3 Algorithm analysis

4.3.3.1 Convergence analysis

Figure 4.5 shows the convergence results of the proposed estimation model. Since the initial upper bound is an infinity (a very large value in numerical computing), the upper bound at the first two iterations are invisible in the figures. The relaxed master problem can be solved very efficiently (in seconds) as it has a limited number of constraints and all constraints have the same format. In addition, we also found that, in our case, subproblems are always feasible if $\mathbf{x}_0$ is feasible, which means all Benders type cuts are optimality cuts, and therefore no feasibility problems needed to solve and no feasibility

cuts are inserted to the master problem. In consequence, these experiments can be solved at a rather cheap computational cost, which indicates the proposed modeling framework has the potential to estimate dynamic OD matrices for large-scale networks. As expected, the algorithm converges within only several iterations for all study intervals, due to the property of the GBD algorithm and the convexity of the problem (GLS estimator). Besides, we found that the convergence of the upper bound can also lead to similar solutions to the ones obtained after the gap between the upper and lower bounds converges. Thus, in practice, one can use solely the upper bound to define the algorithm termination criterion. Considering that the lower bound may sometimes frustrate (which we observed in some experiments), this criterion can be a useful substitute.



(a) 7 am - 8 am **(b)** 8 am - 9 am **(c)** 9 am - 10 am

Figure 4.5: Convergence performance of the GBD algorithm.

4.3.3.2 Estimation quality evaluation

Figure 4.6 illustrates the quality of estimation with respect to the check-in patterns by comparing the empirical and estimated check-in pattern using scatter plots. We can see that all data points are aligned close to the "$y = x$" line in the three experiments. It means the proposed model is capable of recreating the check-in patterns. Recall that the check-in pattern is defined as the vector of the difference in check-in counts between successive time stamps. The good fitting of check-in patterns implies our model can also be used to predict the check-in statistics of activities (venue types).

Similarly, Figure 4.7 visualizes the quality of the estimated OD flows by comparing it with the target OD flows using scatter plots. The R^2 of estimated versus target OD flows in three experiments are 0.91, 0.89 and 0.93, respectively. Overall, the model reaches an acceptable estimate with a slight underestimation in the high-demand OD pairs. We note that the prior OD estimates underestimate the target values by 30% on average. While our model still exhibits underestimation (especially in the high-demand OD pairs), it can, to a certain extent, improve the situation, attributed to the inclusion of $f_p(\cdot)$ and the batch of check-in pattern scenarios. However, the accuracy still cannot meet the requirements of some transportation applications. As can be seen from Figure 4.7, most OD pairs observe a demand of fewer than 100 with many extremely small values. These OD pairs, on one hand, impede the convergence of the algorithm, and on the other reduce

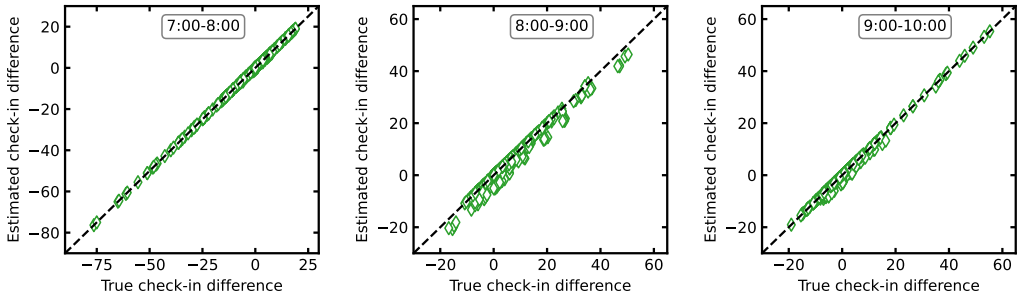


Figure 4.6: Comparison of check-in patterns (performance of the second-stage problem).

the accuracy of the GLS estimator. Thus, three methods can be used to improve the OD flow estimation, including (i) removing the OD pairs with small demand from the model, (ii) applying a reliable weighting matrix in the GLS estimator, instead of using diag(**1**) like in this study, and (iii) replacing the GLS estimator with other estimators, such as maximum entropy.

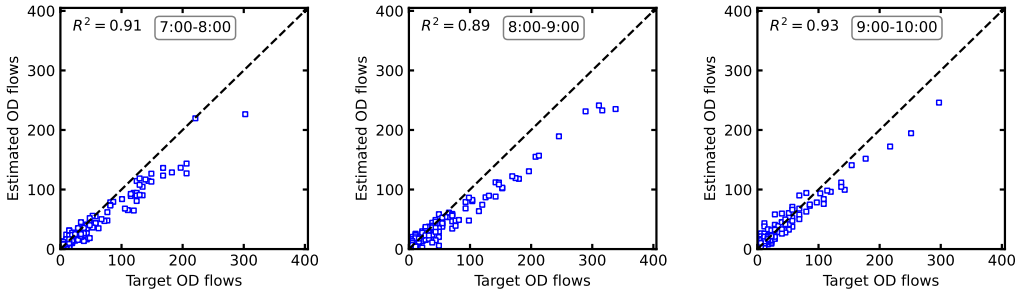


Figure 4.7: Comparison of OD flows (performance of the first-stage problem).

Further, Figure 4.8 compares the theoretic activity shares and the estimated activity shares. Due to the large number of points, we add the heatmap effect in the figure to represent the density of points. Brighter colors mean greater density and vice versa. Overall, there are more points in the range of smaller values. Furthermore, in the figure, we can observe heteroscedasticity among the data points, indicating that the variance of the estimated activity shares increases with higher values of the true activity share. This phenomenon arises due to the activity share constraints expressed by Equation (4.14). These constraints restrict the feasible space of activity flows to be in proportion to the activity share matrix extracted from the activity chain information. As a consequence, the estimation of activity shares is affected by this constraint, leading to varying levels of variability in the data points based on the true activity shares.

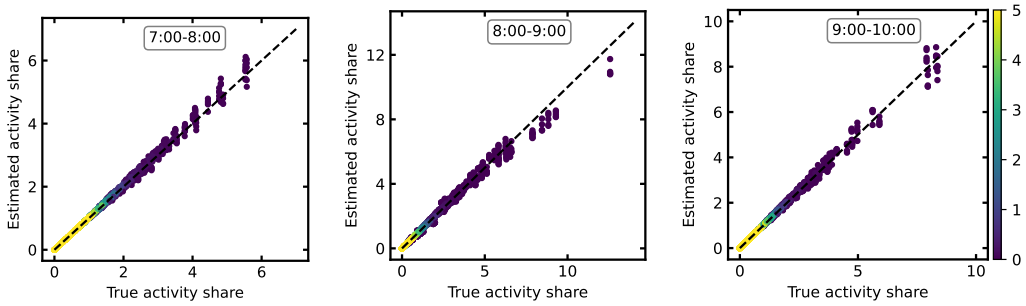


Figure 4.8: Comparison of activity shares (constraint violation).

4.3.3.3 Sensitivity against the number of second-stage scenarios

The number of scenarios is critical to the effectiveness of the method and the improvement in computational cost (compared to simulation-based approaches). Here, we apply four values of N_s (5, 10, 50, and 100) to examine the model and plot their results in Figure 4.9. Four important variables, i.e., OD flows, zonal production, zonal attraction, and check-in patterns, are included in the comparison, and the comparison is conducted with R^2 of the estimated versus target values (with the intercept fixed at 0). Among others, OD flows and check-in patterns are the decision variables of the model for capturing the demand pattern, while zonal production and zonal attraction are counted in the objective function with the purpose of increasing the accuracy of the demand level. For each N_s, considering the randomness of check-in pattern sampling, results from 10 replications are used to create the box plots.

Clearly, $N_s = 10$ outperforms $N_s = 5$ in the estimation of all variables, represented by better average R^2s. It implies that 5 check-in scenarios are insufficient to effectively restrict the search space of the OD matrix, while the new constraints introduced by adding new scenarios can still make a difference. On the contrary, considering too many check-in scenarios also has the possibility of resulting in sub-optimal solutions, especially because of the inclusion of biased scenarios, and adding more complexity. This is evidenced by the worse performance (except for attraction estimation) of the $N_s = 50$ model compared to that of $N_s = 10$. $N_s = 50$ even leads to a smaller R^2 than that of $N_s = 5$ in terms of OD flow and check-in pattern estimation. On the other hand, the $N_s = 50$ model outperforms the models with fewer scenarios in terms of zonal attraction estimation. However, we found that increasing N_s from 50 to 100 would reduce the model performance in all four indicators. This implies that adding too many check-in scenarios into consideration can no longer improve the model performance and might even lead to a performance decline.

Furthermore, we note that the deviation of R^2 in $N_s = 10$ is greater than the other scenarios. It reflects the significance of the quality of check-in scenarios in the proposed model. Despite variations in performance among the four model scenarios, all of them

can output estimations of sufficient quality, thereby indicating the robustness of the proposed approach.

The experiments and comparisons presented here were conducted in an objective state without the presence of zonal attraction ($\{w_x, w_p, w_a, w_s\} = \{1, 10, 0, 1\}$). Thus, zonal attraction serves as a validation metric to evaluate the effectiveness of the model and algorithm. The good fit of zonal attraction in all four scenarios indicates the soundness of the approach. Notably, the zonal attraction scores even outperform those of zonal production, further reinforcing the validity of the model and algorithm.

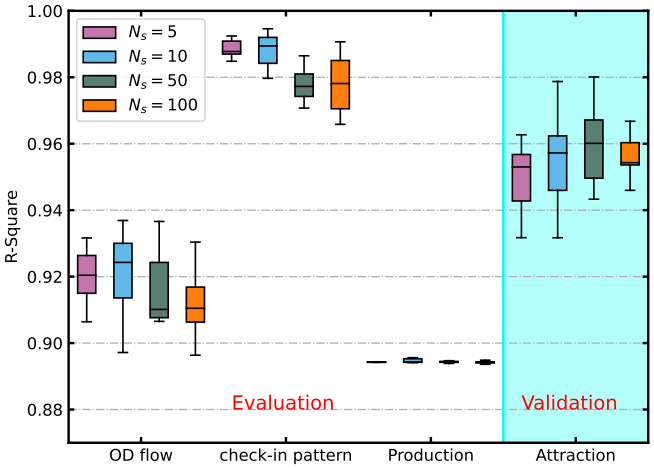


Figure 4.9: Boxplot for R-squares with different numbers of second-stage scenarios.

4.3.4 Objective function analysis

Referring back to Section 4.2.2, the objective function consists of three main types of error terms, i.e., $\mathbb{E}_\xi$, $\{f_p, f_a\}$, and f_x, where $\mathbb{E}_\xi$ is the base model formulated as the two-stage problem, while f_p and f_a can be triggered to add constraints via the trip generation/attraction relationship with check-in patterns subject to the city data, and f_x is a conventional error term widely used in DODE literature to keep the estimated OD patterns close to the prior OD flows. Therefore, the objective function can take multiple states depending on the availability and quality of the data. It is important to analyze the performance of the proposed model under all possible combinations. In this section, we compare the model performance under five different objective function states. Table 4.2 lists the weights of four error terms of these objective states. More specifically, WS tests the model without any restrictions on the OD flows and zonal production and attraction but solely with the check-in pattern optimization. WX0 is used to examine the importance of the prior OD flows. WP0 and WA0 check the necessity of restricting zonal production and attraction, respectively. Finally, state ALL represents the complete form of the objective function.

Table 4.2: Weights of error terms in different objective function states.

State	w_x	w_p	w_a	w_s
WS	0	0	0	1
WX0	0	5	5	1
WP0	1	0	5	1
WA0	1	5	0	1
ALL	1	5	5	1

Figure 4.10 compares the model performance under these objective function states in terms of zonal production and attraction and OD flows. WS obtains a very biased estimation in all variables. This verifies the necessity and reliability of the proposed two-stage stochastic programming framework from the opposite side since if only considering the expectation term in the objective function the model would degrade to a purely single-stage problem. It means despite check-in patterns to a certain extent can be an information carrier of OD patterns, the auxiliary information about the demand level is required for a satisfactory estimation. Adding restrictions on zonal production and attraction as in WX0 can significantly improve the situation. In addition to a good fit in zonal production and attraction, WX0 also provides a fair estimation of OD flows even without feeding any information about OD patterns. Surprisingly, WP0 performs nicely in all three estimation tasks even only having penalties for posterior OD estimates and zonal attraction estimates. Nonetheless, it is also understandable with the consideration that the OD matrix is the mapping result between zonal production and attraction so determining the OD matrix and zonal attraction is capable of reproducing the zonal production as well. Similarly, we can observe comparable results in WA0. Furthermore, as asserted in Section 4.3.3.3, the error terms that are excluded from the objective function, such as zonal production in WP0 and zonal attraction in WA0, can be employed for validation purposes.

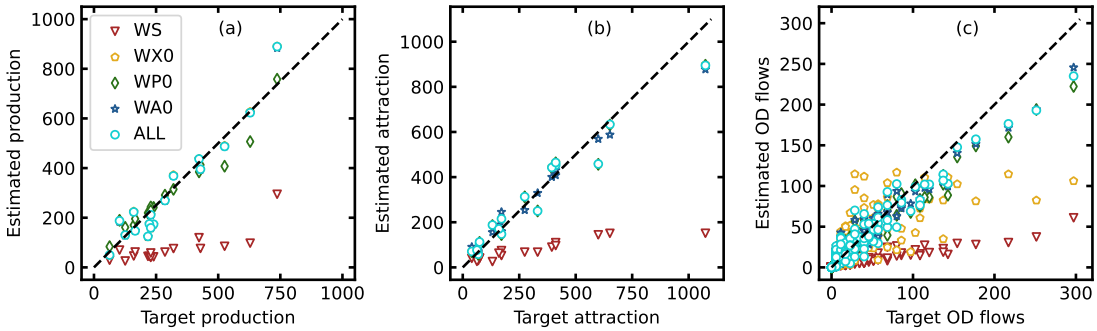


Figure 4.10: Comparison of different objective formulations.

Overall, this proves the superiority of the proposed model in the estimation tasks of zonal production and attraction. It also implies that with further integration of trip distribution information (e.g., the observed trip length frequency distributions from

LSBN activity data), the model can robustly find correct OD matrix patterns, purely relying on LBSN data, without any prior OD estimate information (part focused for future research). In fact, such potential has been presented in WX0 in which only production and attraction data are available. As expected, ALL also leads to promising results. However, its performance cannot explicitly exceed that of WP0 and WA0.

Figure 4.11 shows a clearer comparison of the model under different objective functions by using R^2 as measurement. There is no surprise that WS performs worst in all three tasks. R^2 is calculated by fitting the corresponding data to a linear regression model constrained so that the intercept must equal 0. Under this constraint, the model may obtain a negative R^2 when the model cannot capture the trend of the data, thereby performing worse than the mean estimate (i.e., $\hat{y}_i = 1/n \sum_i^n y_i, \forall i$). This happens in WS, where the fits of production and attraction result in a R^2 of -1.10 and -0.45, respectively. On the contrary, the rest models achieve a score greater than 0.9 in all tasks, except WX0 gets around 0.6 in OD flow estimation. These results conform to the scatter plots in Figure 4.10.

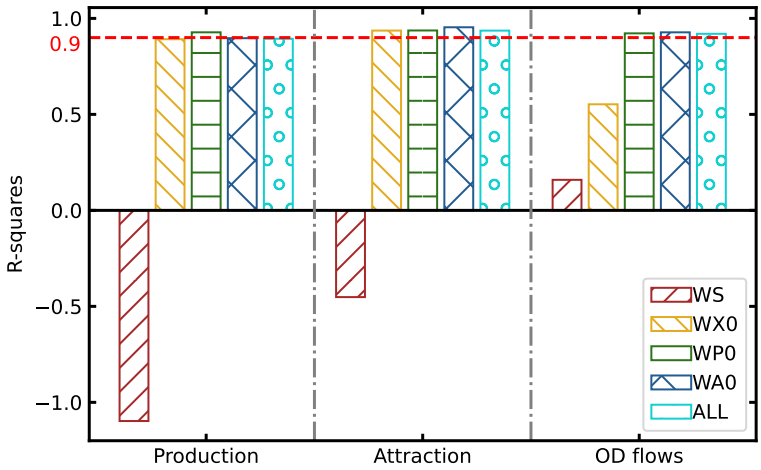


Figure 4.11: R^2 of zonal production, attraction and OD flows under different objectives.

4.4 Conclusions and future work

Considering the stochastic nature of human behaviors and transportation systems, we propose a dynamic OD estimator by utilizing the scenario-based two-stage stochastic programming framework, which integrates the activity chains extracted from LBSN check-in data to model activity-level mobility flows. Given that OD flows are the results of aggregated activity flows, the OD matrix can thus be derived from activity flows easily. More specifically, utilizing the framework, our approach aims to minimize the errors introduced by the inter-zone OD flows and the expected errors of the check-in patterns in the first stage, while minimizing the errors produced by the check-in pattern scenarios in the second stage. The MCMC sampling is implemented to generate a batch

of statistically significant check-in scenarios. Ultimately, a GBD algorithm is applied to solve the two-stage stochastic programming problem, in which the optimal solution is obtained by solving a relaxed master problem and a series of subproblems (one per scenario) alternately.

To evaluate the effectiveness of the proposed approach, we conduct the case study in Tokyo city, Japan, and employ the GLS estimator to measure model performance in estimations of check-in and OD patterns, which on the other guarantees the problem convexity. The experiment results show that the convergence of the GBD algorithm can be attained within several iterations. More importantly, in regard to estimation quality, the model also observes a good fit for check-in patterns, zonal production, zonal attraction, and OD flows. The robustness with respect to the check-in scenarios is also examined and evidenced by incorporating different numbers of check-in scenarios in the second-stage problem. We also explore the objective function formulation space by adopting multiple objective function states with different combinations of error terms. We find that the interdependence of zonal production, zonal attraction, and OD matrix makes the inclusion of all error terms in the objective unnecessary. In particular, considering two of them can already produce estimations as good as considering all. It intuitively follows that the variable excluded from the objective function can serve as the validation metric.

However, the proposed methodology also exhibits the following limitations: (i) In cases of LBSN data lacking venue category information, the methodology may not be viable due to the inherent challenge of differentiating between POIs situated within the same building. (ii) Furthermore, the variability in the representativeness of LBSN check-in data across diverse activities, stemming from potential sampling bias, might affect the reliability of the estimation results.

Chapter 5

Estimating demand pattern changes under disruptions with crowdsensed data

The content of this chapter has been partially presented in the following work:

Lu, Q.-L., Mahajan, V., Lyu, C., and Antoniou, C. (2024a). Analyzing the impact of fare-free public transport policies on crowding patterns at stations using crowdsensing data. *Transportation Research Part A: Policy and Practice*, 179:103944.

5.1 Introduction and research contributions

The proliferation of smartphones equipped with positioning technologies, such as GPS and BeiDou Navigation Satellite System (BDS), has provided a novel means of collecting mobility data, opening up new avenues for demand pattern studies in a more cost-effective manner. These devices generate vast volumes of real-time data on individuals' activities and mobility through location-based services, social networks, and other mobile applications. Crowdsourced data, among others, containing no sensitive personal information, has become a popular source of data in urban mobility modeling including demand pattern analysis. Moreover, their real-time nature enables the evaluation of demand pattern changes caused by disruptions.

This chapter aims to examine the value of crowdsensed data, a typical crowdsourced data, in modeling demand pattern changes at PT stations due to demand disruptions. In particular, we investigate the response of PT stations to the demand surge resulting from fully or partially Fare-Free Public Transport (FFPT), manifested as changes in crowding patterns. FFPT policies allow individuals to utilize PT services free of charge or with reduced financial charges. They serve as a PT incentive to increase PT mode share (Thøgersen, 2009; Cools et al., 2016), which can result in demand-side disruptions if the transit supply is not augmented concurrently.

When analyzing demand patterns at the resolution of PT stations, their heterogeneity in terms of spatiotemporal factors of land use and transportation must be accounted for (Fan et al., 2022). The complex relationship among surrounding land use, transport network, population, temporal features, and historical demand patterns can lead to different crowding patterns. Therefore, the methodology to analyze the impacts of PT pricing policies should consider such heterogeneity. This will contribute to the knowledge of urban mobility and activity patterns and assist policymakers and practitioners in

formulating more rational measures for Public Transport Demand Management (PTDM) (Vongvanich et al., 2023). The proposed methodology can also be applied to predict the disrupted travel demand by integrating with the absolute population information as presented in Vongvanich et al. (2023).

In this chapter, we present a three-step busyness-based framework to empirically evaluate the potential of a wide-coverage crowdsensed dataset in measuring the demand pattern changes at PT stations resulting from demand-side disruptions due to FFPT. Specifically, we first devise a similarity measure based on the changes in crowding patterns at different policy implementation stages. We then integrate a clustering model to identify different types of PT stations based on their crowding patterns and a classification model to infer the relationship between the policy impact and station characteristics.

This chapter offers the following contributions:

- We develop a methodological framework leveraging fine-resolution and wide-coverage PT station busyness data for demand pattern analysis.
- We apply our framework to the case of the 9-Euro ticket experiment in Germany using opportunistic data and provide empirical findings.

The remainder of this chapter is structured as follows. We introduce the three-step framework in Section 5.2. Section 5.3 presents the case studies and experiment results. Finally, major findings, limitations, and future research direction are concluded in Section 5.4.

5.2 Methodology

In this section, we propose a framework driven by crowdsensed check-in data (i.e., busyness data) that can be used to evaluate the impact of FFPT policies on crowding patterns in PT stations and estimate the relationship of the impact with station characteristics.

Specifically, we develop a three-step busyness-based evaluation framework, as shown in Figure 5.1. In addition to the collection and preprocessing of crowdsensed data and station attributes, the framework integrates the following three components:

Step (1) A histogram-based feature engineering approach is adopted to encode the crowding pattern changes in PT stations. The histogram method is a means for estimating the probability density distribution of a certain random variable. It can provide an overview of the impact of the policy of interest on the concerned stations. Besides, factor analysis is applied to extract significant patterns from raw station features and reduce noise and feature dimensions.

Step (2) A clustering model is used to identify and label different types of PT stations based on the changes in their crowding patterns. This step is to mine and characterize the changing patterns of PT station crowding patterns. We define these clusters as busyness-based station types (i.e., Y_c in Figure 5.1), which will serve as class labels inputted to the classification model in the following classification step.

Step (3) Finally, a classification model is trained on the spatiotemporal station attributes and the busyness-based station types identified by the clustering model in Step (2) so as to infer the association between them.

Furthermore, it is worth mentioning that in most platforms providing crowdsensed services, the busyness value of a POI may be unavailable if there is a limited number of users opting for location services. In order to estimate the crowding pattern changes, it is necessary to impute the missing data contained in the raw busyness dataset beforehand.

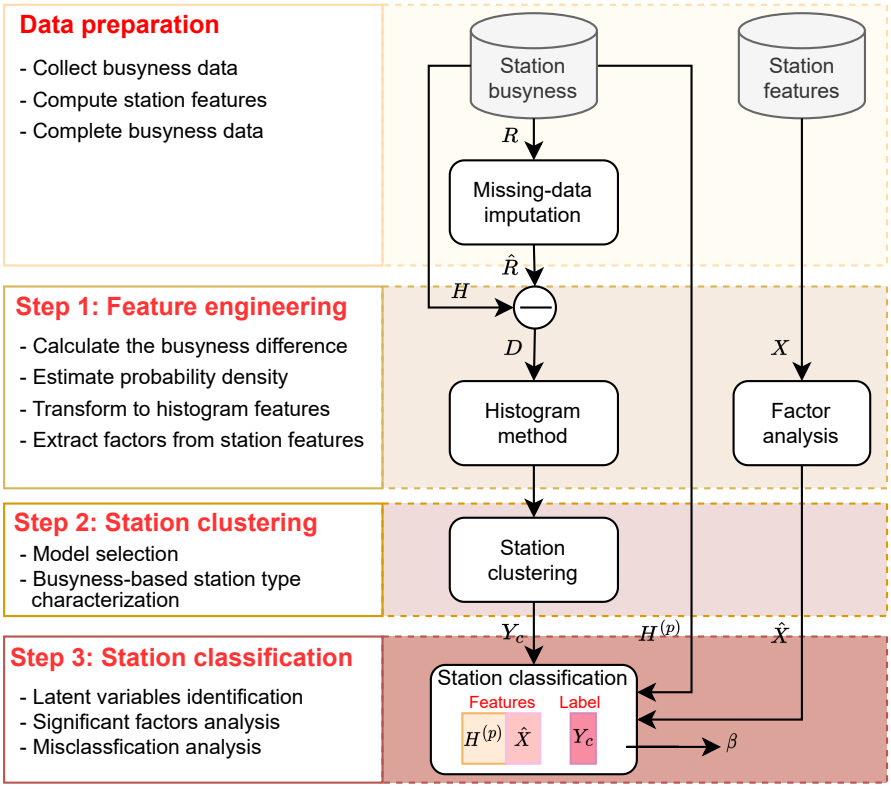


Figure 5.1: Three-step busyness-based evaluation framework for public transport policies.

5.2.1 Histogram-based feature engineering

Generally, for a specific time stamp, crowdsensed services provide two busyness values for POIs, namely, historical and live busyness. The deviations of live busyness from the historical ones can be used as a proxy for PT demand pattern changes. Based on the histogram method, we use the busyness deviations to construct a measurement to describe the similarity between PT stations in terms of the impact of fare-free policies. Figure 5.2 provides a graphical illustration of the proposed histogram method.

Let $\boldsymbol{H} \in \mathbb{R}^{n_s \times n_t}$, $\hat{\boldsymbol{R}} \in \mathbb{R}^{n_s \times n_t}$ denote the historical and imputed live busyness matrices, where n_s and n_t denote the number of stations and time intervals respectively. We

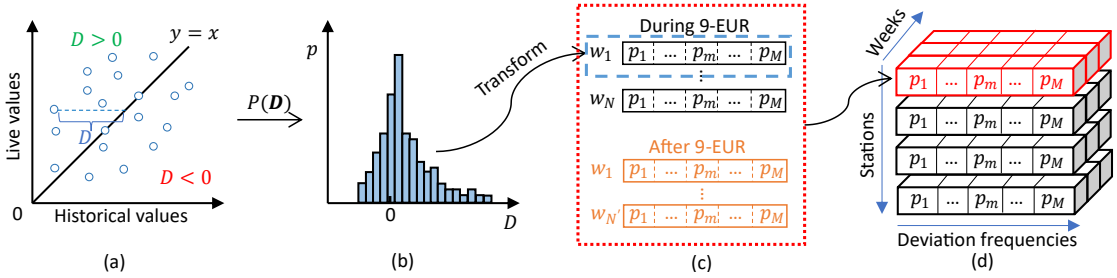


Figure 5.2: Feature engineering based on the histogram method.

denote H_{ij} and $\hat{R}_{ij}$ the historical and imputed live busyness of station i at time interval j, which can be absolute (number of visitors) or relative values (within a range like $[0, 100]$) depending on the data source. For instance, Foursquare provides the number of visitors, while Google Maps only provides the normalized value in the range $[0, 100]$. The busyness deviation matrix can then be calculated as $\boldsymbol{D} = \hat{\boldsymbol{R}} - \boldsymbol{H}$ (Figure 5.2a). The histogram method can then be applied to approximate the probability distribution of the respective busyness deviations for each station and each week separately. As shown in Figure 5.2b, for one week of data of a station, we can plot the deviations using a histogram; there are 168 deviation values within a week for a station at which data are updated hourly.

In the histogram method, the sample space is divided into M non-overlapping bins, each with width Δ_m. We then count the deviation values located in each bin K_m and approximate the density function by

$$p_m = \frac{K_m l_t}{168 \Delta_m}, \quad \forall 1 \leq m \leq M \tag{5.1}$$

where l_t (h) is the length of each time interval, and 168 is the number of hours each week. Repeating the calculation for every station and every week (Figure 5.2c), we can obtain a probability density tensor of busyness deviations with dimension $n_s \times M \times n_w$ where n_w is the number of weeks (Figure 5.2d). This feature engineering method extracts representative features that can reflect the crowding pattern changes from the PT station busyness data.

5.2.2 Clustering for labeling busyness pattern changes

Denote $\boldsymbol{Q} \in \mathbb{R}^{n_s \times n_p}$ (n_p is the total number of histogram bins of all weeks under consideration) as the probability density matrix formed by stacking probability density vectors of different time intervals and stations along two dimensions. The Gaussian Mixture Model (GMM) is employed to perform the station clustering task. Given the training dataset and a GMM configuration (w.r.t. number of Gaussian components, type of covariance matrices), model parameters (i.e., mean vectors, covariance matrices, and mixture weights) are estimated to maximize the GMM likelihood. Note that each component density is a n_p-variate Gaussian function . The problem can be addressed via

the expectation-maximization algorithm proposed in Reynolds et al. (2009). In terms of hyper-parameters, while they are often determined by the amount of data, one can also make a decision based on the Bayesian Information Criterion (BIC) or the Akaike Information Criterion (AIC). In statistics, these two criteria are used for model selection, particularly to prevent from the overfitting issue caused by adding too many parameters into the model. The present clustering model reveals the heterogeneous impacts of fare-free policies among PT stations and provides labels to the classification model introduced in the following.

5.2.3 Station classification using spatiotemporal attributes

Many factors can lead to the difference in the response of PT stations to PT policies, including static features such as the network structure, and dynamic features such as activity patterns in the station vicinity. Table 5.1 summarizes some typical influencing factors together with their representative features. While the clustering model can identify different changing patterns of crowding patterns at PT stations, we move one step forward by training a classification model to identify the significantly correlated factors with the changing patterns. In this chapter, we primarily consider the four categories of features listed in Table 5.1.

Table 5.1: Features considered in the classification model.

Feature categories	Features
Location & Context	Dummy variable for cities and station types (U-Bahn, S-Bahn)
Nearby activity	Counts of different types of POIs around the station, population
Crowding pattern	The crowding pattern of the station for one week (before policy implementation)
Network structure	Features used in complex network analysis: # edges, # nodes, average node degree, etc.

"Location & Context" features are able to capture the local economic situation, communities' life culture, and development pattern to some extent (Fan et al., 2022). U-Bahn and S-Bahn in Germany refer to urban subway and city rapid railway, respectively. Other categorizations in the context of other countries are also applicable. Similarly, "Nearby activity" features represent the land use, demand production and attraction of the adjacent area around the PT station. Apart from the statistics of POIs, the busyness data of POIs are also valuable for modeling neighborhood activity patterns. "Crowding pattern" features provide information related to the demand pattern of the respective PT stations under normal operation conditions. They are extracted from the historical busyness data collected before implementing new pricing measures. Furthermore, "Network structure" features can capture the traffic conditions of local private transport, which are very important in PT-related assessments given the strong inter-correlation between private and public transport.

It is worth pointing out that "Nearby activity" features might be highly correlated. It follows that dimension reduction techniques can be applied to reduce the number of features and preserve most information contained in the original features at the same time. This is also applicable to the "Crowding pattern" features. In this study, we apply factor analysis to achieve this. Factor analysis also serves as an exploratory analytical instrument, and its application is deemed justifiable to the extent that the resulting factors can be meaningfully interpreted.

We apply the LightGBM to perform the classification task. An introduction of this model can be found in Section 3.3.2.2. A sequential feature selection is used to find the best combination of features for classification. A 10-fold cross-validation is also conducted to evaluate the performance of the LightGBM classifier on the busyness-based station types classification task.

5.3 Case study and results

In 2022, the German federal government introduced the so-called "9-EUR Ticket" in response to the escalating fuel and energy costs resulting from the geopolitical crisis in Ukraine (Loder et al., 2023). The primary objective of the measure was to encourage commuters to shift from private vehicles to PT by providing them with a monthly flat rate of 9 EUR, which enabled them to travel on all regional, local, and urban PT services, except for long-distance passenger services such as Intercity Express, from June 1 to August 31, 2022. It provides an unprecedented opportunity for understanding the impact of ticket schemes with substantial discounts on PT operations and the potential of crowdsensed data for capturing demand pattern changes. To this end, GPT data are collected for more than 2000 PT stations in Germany to drive the busyness-based framework. The introduction of GPT data can be found in Section 3.4.1.3 and Section 3.4.1.4.

5.3.1 Features and data analysis

The data collected before 9-EUR are used to extract the "Crowding pattern" features for classification, while the data collected during and after the 9-EUR period are used in the histogram method for clustering to identify changing patterns.

Regarding the other features adopted in the classification model, the population is estimated using the population density map of Germany in 2019 (HDX, 2023). For the nearest 5 and 10 stations, we calculate the mean and standard deviation of the population densities to capture the population characteristics. Network neighbors of a given station are defined as the closest stations to it in the entire PT network. These features capture the connectivity characteristics regarding the busyness level within the adjacent area of the neighboring stations. POI statistics are collected from OpenStreetMaps (84 types are reserved after a POI counts filter). Both the population and POI statistics are estimated and counted within a square are with a side length of 1 km. "Network structure" features are computed using the Python package OSMnx (Boeing, 2017).

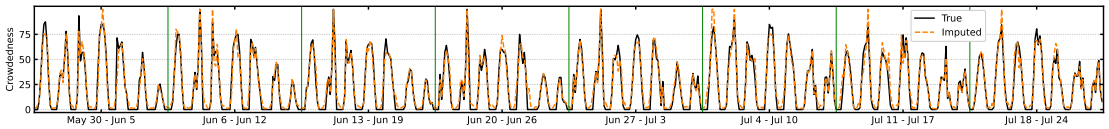
In terms of historical average patterns, two demand peaks are observed on weekdays, including a morning peak at 8 AM and an afternoon peak at 3 PM, while only one

peak shows up on Saturdays and Sundays at around 3 PM. It means the daily crowding patterns are different between weekdays and weekends. Furthermore, the busyness of stations is higher on weekdays than on weekends during the daytime. In contrast, there is a higher busyness level after 8 PM on Saturday, possibly due to more nightlife activities on Saturday evenings. Regarding live average patterns, an apparent lift shows up in Saturday's pattern and Sunday's pattern in all time intervals, compared to the historical ones. It means the implementation of 9-EUR significantly stimulates travel on the weekends. On the contrary, on the weekdays, while the live pattern keeps at a similar level to the historical pattern during the daytime, it presents a higher popularity in the evening. Another interesting point is the standard deviations of live patterns across different stations are much greater than those of historical patterns, implying that the influence of 9-EUR has a nature of inter-station variability. This lends support to the development of the three-step framework and the necessity of integrating station clustering to accurately uncover the policy's influence.

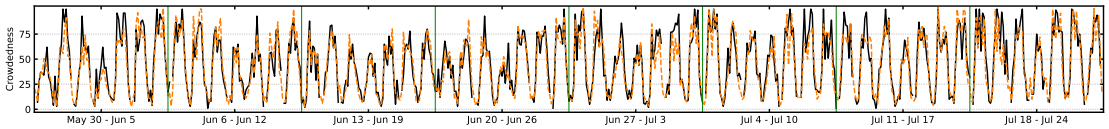
5.3.2 Crowdsensed data imputation

Further, we notice the common missing value issue of live busyness in GPT data; when there are insufficient visitors, the live busyness may not be available. Therefore, it is necessary to apply a data imputation method to complete the live busyness dataset.

Due to the potential difference in busyness patterns during collection periods during and after 9-EUR, live data imputation is performed separately for the two periods. The missing rates during and after 9-EUR are 21% and 33%, respectively. Considering the imbalance issue of missing patterns in the dataset, i.e., some stations observe many missing values while observing none otherwise, we apply the Probabilistic Matrix Factorization (PMF) (Mnih and Salakhutdinov, 2007) to address this problem. By PMF, the live busyness matrix will be modeled as a product of two lower-rank matrices, i.e., latent station and time feature matrices. The missing entries in the matrix will be imputed by maximizing the log of a posteriori distribution over station and time features.



(a) Imputation result of historical data



(b) Imputation result of live data

Figure 5.3: Comparison of true and imputed popularity (Jaczostr. station, Berlin).

To evaluate the performance of the imputation method, we artificially design two random missing scenarios with the same missing rates using the respective historical datasets. In these two synthetic scenarios, PMF leads to mean absolute percentage errors (MAPEs) of 2.3% and 1.0%, respectively. As an example, Figure 5.3 shows the busyness time series of the Jaczostr. station, Berlin, in the first eight weeks of 9-EUR, together with their imputed values. We can see that despite live data fluctuating more than historical data, it also follows a clear daily pattern and the imputed values fit quite well with the ground truth.

5.3.3 PT station clusters characterization and analysis

In the histogram method, we set all bins to the same width $\Delta_m = 10, \forall m$. Since the busyness deviation will be in the range of $[-100, 100]$, $M = 20$ bins are specified in histograms. According to Equation 5.1, after specifying the same l_t and Δ_m for all bins, the probability density is proportional to the counts of deviation values, i.e., $p_m \propto K_m, \forall m$. Thus, we use K_m rather than p_m in the following experiments for a more intuitive understanding. Additionally, the variables representing the range of $[-100, -60)$ and $[80, 100]$ are removed due to their limited capability of distinguishing station samples, i.e., almost all stations have the same value. Figure 5.4 compares the performance of GMM models with different hyper-parameters with respect to the number of components and the type of covariance matrices. A grid search is performed to find the best hyper-parameters based on BIC. A lower BIC is preferable. The top subplot presents the results of considering all histogram variables, while the bottom subplot presents that of after removing the variables representing $[-100, -60)$ and $[80, 100]$. BIC is used to avoid overfitting by incorporating a penalty term for the number of parameters in the model, which interprets why the GMM models with full or tied covariance matrices (non-empty values in all elements) for each component perform worse than the other two types of GMMs. Furthermore, we can see that removing the variables incapable of distinguishing stations can improve the robustness of the GMM with diagonal covariance to the number of components. Finally, focusing on the bottom case, although the GMM models with spherical and diagonal covariance matrices show similar performance regardless of the number of components, the GMM with 3 components and diagonal covariance matrices leading to the smallest BIC is selected for interpretation purposes.

After reordering, the three categories of stations are indicated in Figure 5.5. The x-axis is the bins indices, with every 14 bins constituting a histogram representing one week of data because we have removed the six low-impact intervals mentioned above. We have 14 weeks for the 9-EUR period and 9 weeks for the post 9-EUR period, which are indicated with different colors in the figure. The y-axis represents different PT stations. Red horizontal dashed lines separate different categories of stations. Zoom-in views of some stations of each category are also provided to better illustrate the difference in their crowding pattern changes. A clear distinction can be observed between every two categories. To be specific, we characterize stations based on the heterogeneous impacts of 9-EUR on their crowding patterns as (i) Cluster 1 – unaffected stations (146 stations). Their demand patterns are almost not affected by the policy. Moreover, their crowding

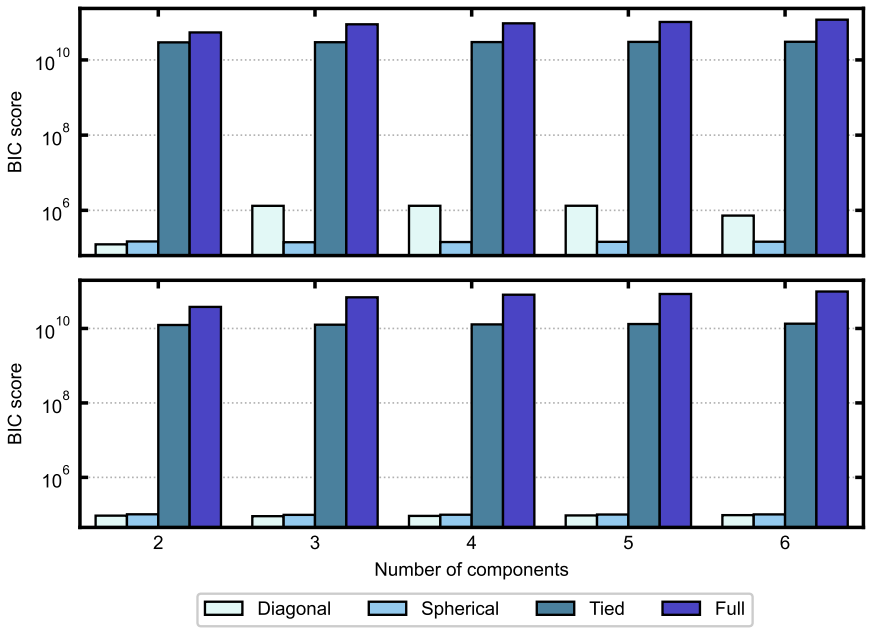


Figure 5.4: GMM models comparison based on BIC. (top) Considering all variables resulting from the histogram method; (bottom) After removing insignificant variables.

pattern deviations are more stochastic than the others; (ii) Cluster 2 – mildly stimulated stations (92 stations). These stations' demand increase during 9-EUR and recover to the original demand level slowly or keeps a similar level afterward; (iii) Cluster 3 – intensely stimulated stations (55 stations). These stations' demand increases significantly during 9-EUR and reduces immediately afterward.

Figure 5.6 provides the GPT time series samples during and after 9-EUR for each station type to better illustrate the heterogeneity among different types of stations. These time series demonstrate similar phenomena to those described above based on histogram-based variables in terms of the changes in crowding patterns, which, on the other hand, reflect the correctness of the characterization of stations.

Figure 5.7 provides the composition of stations in five major cities in Germany that have more than 20 stations in the dataset. The cities are ordered based on the number of "unaffected" stations (cluster 1). We can see that the composition is very different from city to city. Specifically, most stations in Stuttgart and Berlin are "unaffected", while more than half of stations in Hamburg and Frankfurt are "mildly stimulated" and "intensely stimulated" respectively. The PT use in Berlin did not change much due to 9-EUR, while that in Frankfurt (a city of the Hessen state) increased significantly. On the other hand, Munich has nearly balanced numbers of "unaffected" and "mildly stimulated" stations. Moreover, 87.3% of "intensely stimulated" stations (55 in total) are from Frankfurt.

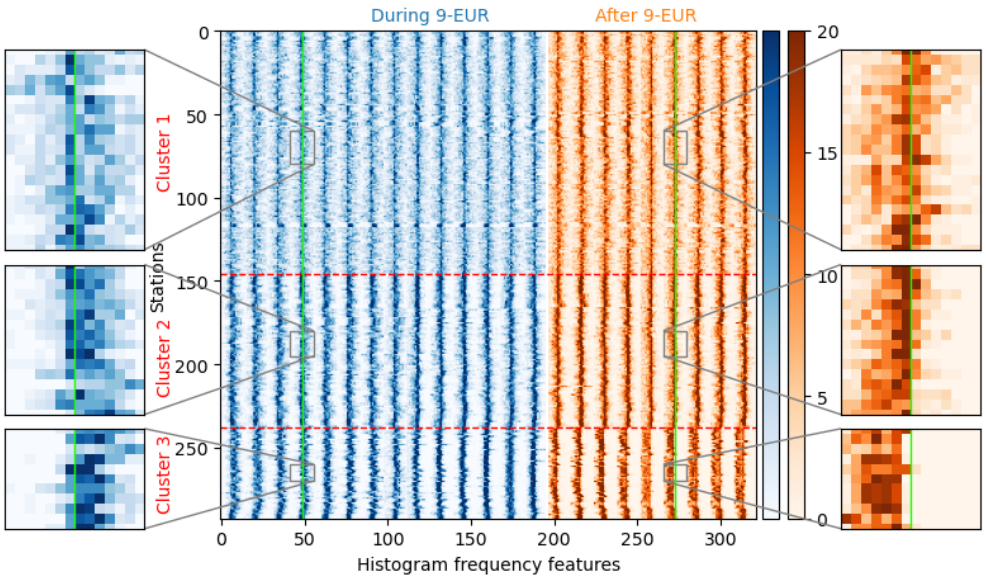


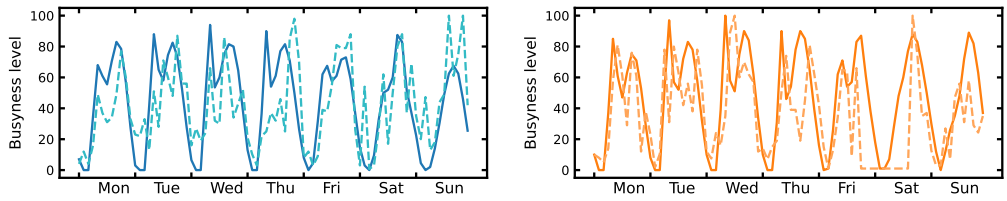
Figure 5.5: Station clusters based on busyness deviations. (blue) During 9-EUR; (orange) After 9-EUR. Green vertical lines indicate the centers of the respective histograms.

5.3.4 Latent variables extraction based on factor analysis

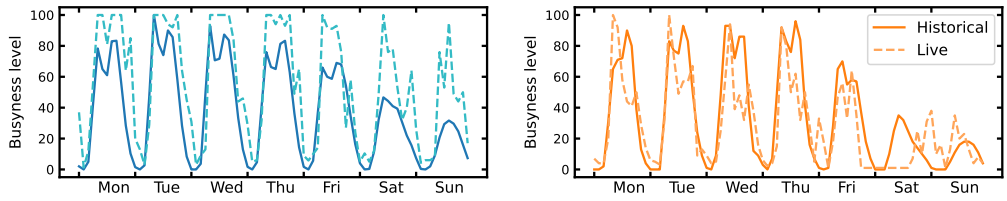
Factor analysis with the VARIMAX (Kaiser, 1958) rotation method is conducted on crowding pattern features to extract significant factors that can describe the demand patterns of PT stations. Only factors that can explain more than 2% of the total variance and have an absolute factor loading greater than 0.5 are considered for characterizing the factor, which leads to 11 factors from 84 crowding features.

Figure 5.8 depicts the factor loadings of features in different factors. The x-axis shows different periods within a week, while the y-axis indicates the factors which satisfy the constraints. Darker colors indicate a large absolute value of the corresponding factor loading. We denote these factors extracted from crowding pattern features by CP0 to CP10. Clear distinctions do exist among them. In particular, according to the factor loadings, we can interpret these factors as representatives for different times of the day. We characterize these factors as listed in Table 5.2. From these factors, we can also see a clear distinction between weekdays, Saturdays, and Sundays. Therefore, according to the time intervals with significant factor loadings, we define three basic types of factors: weekday-based, Saturday-based, and Sunday-based factors. All factors can be categorized under at least one category of them. Then, the factors belonging to both Saturday-based and Sunday-based Saturday- and Sunday-based factors would be defined as weekend-based factors, while those relating to all days of the week would be daily factors.

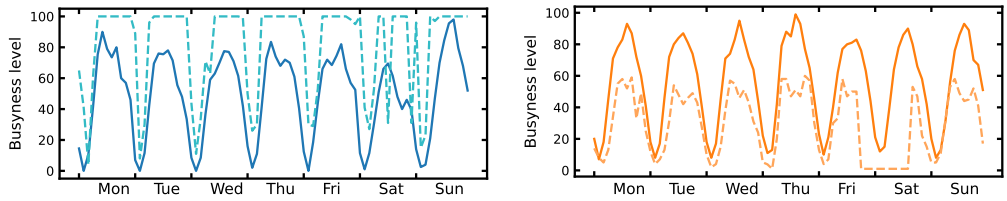
Essentially, these factors represent the demand patterns at different periods of the week and are latent variables contained in the crowding pattern features. For instance, CP0, associated with the period 14:00 - 22:00, represents the demand pattern at PT stations



(a) Cluster 1 during 9-EUR (Jun 20 - Jun 26) **(b)** Cluster 1 after 9-EUR (Oct 10 - Oct 16)



(c) Cluster 2 during 9-EUR (Jun 20 - Jun 26) **(d)** Cluster 2 after 9-EUR (Oct 10 - Oct 16)



(e) Cluster 3 during 9-EUR (Jun 20 - Jun 26) **(f)** Cluster 3 after 9-EUR (Oct 10 - Oct 16)

Figure 5.6: GPT time series samples of the identified busyness-based station types.

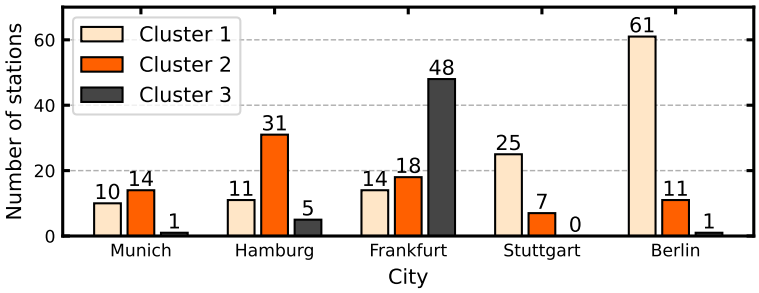


Figure 5.7: Station types distribution in major cities ($n_s \geq 20$).

in the afternoons and evenings. The negative loadings in these intervals say that the busyness level at PT stations during this time period is related in the opposite direction from the factor. The variance each factor can explain is also provided in Table 5.2.

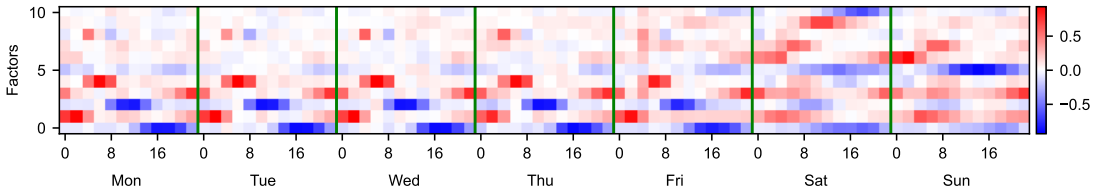


Figure 5.8: Factors extracted from crowding pattern features.

Table 5.2: Factors extracted from crowding pattern features.

Factor	Type	Time	Interpretation	Loadings
CP0	Daily	14-22	Afternoon and evening	Negative
CP1	Daily	0-6	Early morning	Positive
CP2	Weekday-based	10-16	Weekday lunchtime	Negative
CP3	Daily	22-2	Mid-night	Positive
CP4	Weekday-based	4-10	Weekday breakfast time	Positive
CP5	Sunday-based	10-22	Sunday after morning	Negative
CP6	Weekend-based	0-6	Weekend early morning	Positive
CP7	Weekend-based	6-10	Weekend breakfast time	Positive
CP8	Weekday-based	4-6	Weekday before breakfast	Positive
CP9	Saturday-based	10-14	Saturday lunchtime	Positive
CP10	Saturday-based	16-22	Saturday evening	Negative

Further, factor analysis is also conducted on the highly correlated (here, features with an absolute correlation greater than 0.6) features in the rest of features. These features all belong to either POI statistics, population densities, or network features. Following the same procedure, we obtain 9 factors. The factor loadings are provided in Figure 5.9. We denote these factors by PPN0 to PPN8 with "PPN" representing "POI statistics", "Population" and "Network" respectively.

We can interpret PPN factors as in Table 5.3. Due to the large amount of POI features and the difficulty of explaining their intrinsic correlation, we do not explicitly categorize different types of POIs and finely interpret the associated latent variables. Instead, we only distinguish the POI-associated factors from the sign of factor loadings, resulting in three POI-based factors with positive loadings (PPN0, PPN3 and PPN5) and two POI-based factors with negative loadings (PPN1 and PPN8). On the contrary, for those factors emphasizing population or network features, we can provide clear interpretations based on the nature of associated features. In particular, the factor analysis leads to two population-based and two network-based latent variables. The two population-based factors indicate population and population variance, respectively, given that the

former is correlated with the population densities of the station vicinity and its network neighbors, while the latter is correlated with the variance in their population densities. The two network-based factors indicate the network summation describing the counts or summations of network elements (nodes, edges, intersections, etc.), and the network average describing the average of network elements. Note that while the factor loadings of the two population-based factors have opposite signs, both network-based factors have positive loadings on the associated features.

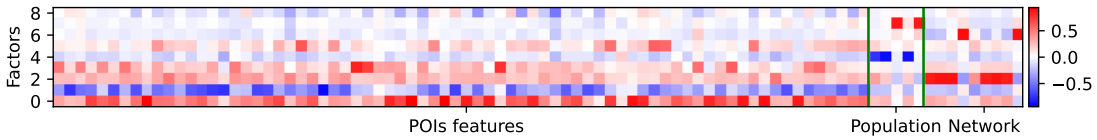


Figure 5.9: Factors extracted from highly correlated features.

Table 5.3: Factors extracted from highly correlated features.

Factor	Interpretation	Loadings
PPN0	POI positive (1)	Positive
PPN1	POI negative (1)	Negative
PPN2	Network summation	Positive
PPN3	POI positive (2)	Positive
PPN4	Population	Negative
PPN5	POI positive (3)	Positive
PPN6	Network average	Positive
PPN7	Population variance	Positive
PPN8	POI negative (2)	Negative

5.3.5 Classification model evaluation

LightGBM is applied to infer the relationship between the selected station features (including the latent variables resulting from factor analysis) and the impact on crowding patterns. Based on the cross-validation, the model achieves a weighted F1 score of 0.704 and a balanced accuracy of 0.701. With the information gain as the importance measurement, the significant features with an information gain greater than 100 are given in Figure 5.10. As can be seen, the importance of the dummy variable indicating whether the station is located in Frankfurt is more than two times of most of the rest of variables. Referring back to Figure 5.7, we found that 60% of Frankfurt's stations (48 out of 80) are "intensely stimulated" stations, composing 87.3% of this station type. It is thus understandable that this variable attains the highest importance. Similarly, the variable indicating whether the station is located in Berlin illustrates a significant importance level since 83.6% of its PT stations belong to "unaffected" stations. Additionally, the

second to the ninth most important variables (8 out of 13) are latent variables identified in Section 5.3.4. Four of them are extracted from crowding pattern features, representing the afternoon and evening pattern (CP0), the pattern before breakfast on weekdays (CP8), the early morning pattern (CP1), and the evening pattern on Saturday (CP10). This means the influence of 9-EUR on a given PT station also has a strong relationship with its daily crowding patterns. The other four factors are PPN factors, representing demand patterns of the POIs near the station (PPN3 and PPN8), population (PPN4) and population variance (PPN7). The "Self loop" variable (i.e., percent of edges that are self-loops) is one "Network structure" variable described in Table 5.1. All feature categories listed in Table 5.1 show up in Figure 5.10.

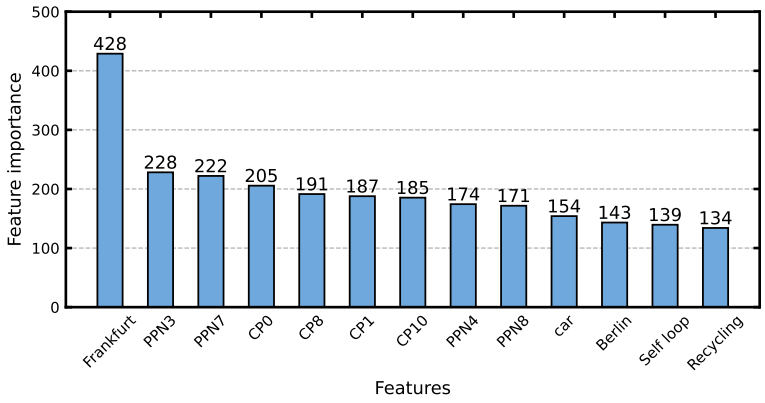


Figure 5.10: The most important features with information gain greater than 100.

5.4 Conclusions and future work

This chapter has presented a three-step busyness-based evaluation framework to empirically analyze the PT demand pattern changes under demand-side disruptions, with a focus on crowding patterns at PT stations. Using the example of the 9-EUR ticket implemented in Germany in 2022, we have showcased the effectiveness of crowdsensed data in estimating the varying sensitivity of PT stations to fare intervention and inferring the factors influencing the difference.

The clustering model has successfully identified three categories of stations based on their crowding pattern changes: (i) Unaffected stations (demand patterns are not affected by the intervention), (ii) Mildly stimulated stations (demand increases during the implementation period and recovers to the original state slowly or maintains a similar demand level afterward), (iii) Intensely stimulated stations (they are very sensitive to the intervention, as shown by the dramatically increasing demand during the implementation period and the sudden reduction afterward). It has revealed that most stations in Stuttgart and Berlin are "unaffected", while more than half of stations in Hamburg and Frankfurt are "mildly stimulated" and "intensely stimulated" respectively. Munich has

balanced numbers of "unaffected" and "mildly stimulated" stations. Further, the classification model suggests that the station location, activity patterns, population variance within the area around the station and around its nearest neighbors in the network, and demand patterns in the afternoon and evening, and weekday early morning play a significant role in the crowding pattern changes of stations.

However, the framework can still be further improved by integrating other relevant features, such as accessibility to the city center. Future research could also consider including POI statistics around the areas of the station's network neighbors, given the significance of their population densities (and variance) in the classifier. Another important limitation of this study is that the crowding patterns of other POIs near the station, which could capture the activity patterns nearby, are not considered and collected. However, these activity patterns are expected to be highly correlated to the crowding patterns since people usually travel to perform activities. Furthermore, while the histogram features extracted from the distribution of crowding pattern deviations can effectively identify PT stations with similar crowding pattern changes, it fails to capture the similarity in the temporal patterns of stations' busyness levels. Future works could explore the development of a method capable of combining both of these aspects.

Chapter 6

Traffic resilience indicators based on macroscopic fundamental diagrams

The content of this chapter has been partially presented in the following work:

Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2024c). Traffic resilience quantification based on macroscopic fundamental diagrams and analysis using topological attributes. *Reliability Engineering & System Safety*, 247:110095.

6.1 Introduction and research contributions

A reliable indicator of system functionality is critical to the evaluation of system resilience. Although the resilience concept has been adapted to transportation systems for more than two decades, there is no universally accepted definition of system functionality in the context of urban road transportation yet. By convention, topological measures based on complex network theory, which can represent the structural properties (e.g., connectivity and accessibility) of the network, are used. However, these measures cannot capture traffic dynamics in the evaluation. On the other hand, trip-based indicators, such as network average travel time (Dingil et al., 2019; Arango et al., 2023), average speed (Hoogendoorn et al., 2015), and demand served (Chen and Miller-Hooks, 2012), have also been adopted to overcome the drawbacks of the topology-based ones. However, these indicators based on direct trip information are usually sensitive to travel demand levels and patterns, and cannot be regarded as a property of the transportation system.

With the definition of the MFD (Geroliminis and Daganzo, 2008) and the success of its applications to perimeter control (e.g., Zhong et al., 2018; Chen et al., 2022), one can construct reliable and representative traffic resilience indicators by relating the MFD dynamics and system functionality. The MFD describes the relationship among the space-mean production, the space mean state, and spatial homogeneity in traffic flow, offering a way to evaluate network performance without requiring detailed traffic physics. More importantly, MFD is an intrinsic property of a homogeneously congested transportation network (Huang et al., 2020; Su et al., 2020). Therefore, compared with other measures, it holds the superiority of integrating information on transportation network characteristics, traffic dynamics, and travel demand patterns. Note that transportation system disruptions always render recurring or non-recurring congestion. This provides an opportunity for measuring system resilience by comparing the MFD dynamics before

and after disruptions. To date, several MFD-based resilience indicators have been proposed, including the link criticality index based on the shape of MFDs (Kim and Yeo, 2017), the performance indicator based on travel production (Amini et al., 2018), and the traffic resilience index based on total congestion deviation (Gao et al., 2022).

As a step in that direction, we present enhanced MFD-based traffic resilience indicators in this chapter to address the defects in its predecessors. To be specific, we theoretically discuss and compare the influencing mechanisms behind hyper-congestion and supply-side disruptions, which are the transportation disruption cases mostly discussed in the literature (e.g., Gao et al., 2022; Zhang et al., 2022a; Pei et al., 2024), highlighting the need to differentiate the calculation of traffic resilience based on the type of disruption. To illustrate the practical value of our approach, we conduct case studies on two real networks, Munich, Germany, and Kyoto, Japan.

This chapter contributes to the existing body of knowledge on traffic resilience and its practical implementation. Specifically,

- We develop novel resilience indicators for transportation systems based on the well-established MFD concept, emphasizing their function in serving traffic flows.
- We validate the robustness and reliability of the proposed indicators through numerous experiment scenarios of two real networks.

The rest of the chapter is structured as follows. Section 6.2.1 provides a brief introduction to basic concepts in the MFD. Section 6.2 presents the MFD-based traffic resilience indicators and discusses the necessity of distinguishing between hyper-congestion and supply-side disruptions. Then, Section 6.3 introduces the case studies and experiment results. Finally, Section 6.4 concludes the chapter with the findings and future directions for research.

6.2 Methodology

In order to address the limitations of existing resilience indicators discussed above, we propose comprehensive MFD-based traffic resilience indicators in this section, partially building upon the work presented by Amini et al. (2018). Our approach acknowledges the need to differentiate between the evaluation of traffic resilience to hyper-congestion and supply-side disruptions due to the distinct mechanisms through which they exert influence on the system. This differentiation is crucial for devising and implementing appropriate preparedness and recovery strategies to mitigate their impact and enhance network resilience. In particular, hyper-congestion typically involves recurring events, allowing for the application of conventional network traffic optimization methods to improve resilience. On the other hand, supply-side disruptions are typically non-recurring events that require event-specific and tailored approaches. By distinguishing between these two types of disruptions, we can better understand the unique challenges they pose and identify the most effective response measures. We introduce the notation listed in Table 6.1 for ease of explanation.

6.2.1 Macroscopic fundamental diagram

The MFD describes the relationship between the number of vehicles on a network and their average speed. Essentially, it provides insights into the network's internal flow (represented by the number of vehicles per roadway length or weighted average density) and its outflow (represented by the trip completion rate or weighted space-mean flow). Figure 6.1 illustrates the typical parabolic shape of an MFD. One of the key advantages of the MFD is its simplicity and applicability in real-time control problems, as it does not require extensive traffic data from multiple sources or complex estimation methods. The MFD can be constructed using raw field data obtained from common loop detectors or floating car data. Moreover, the MFD is an intrinsic property of a homogeneously congested transportation network and remains independent of traffic demand. This makes it a valuable tool for analyzing and optimizing transportation systems under varying conditions.

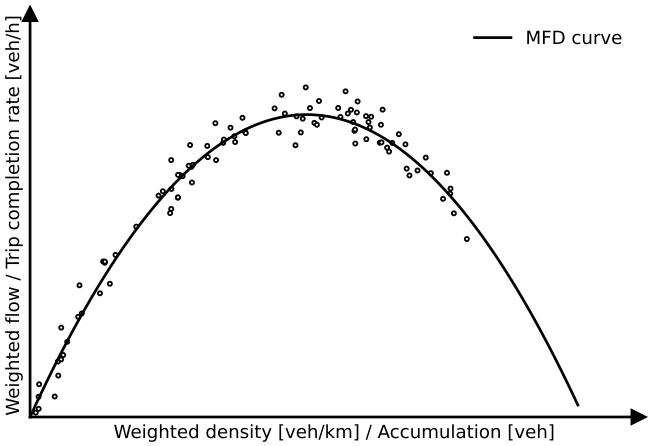


Figure 6.1: An example of the macroscopic fundamental diagram.

Specifically, for a certain time step, the weighted average density and the weighted space-mean flow are given by

$$k(t) = \frac{\sum_{i\in\mathbb{L}} l_i k_i(t)}{\sum_{i\in\mathbb{L}} l_i} \tag{6.1}$$

$$q(t) = \frac{\sum_{i\in\mathbb{L}} l_i q_i(t)}{\sum_{i\in\mathbb{L}} l_i} \tag{6.2}$$

where $\mathbb{L}$ is the set of links with loop detectors installed, l_i presents the length of link i, k_i and q_i represents the vehicle density and flow at link i, respectively. Meanwhile, accumulation is the total number of vehicles in the network, which is usually difficult to acquire in practice. Nonetheless, since the roadway length is fixed, the accumulation is proportional to the density. It can be estimated as

$$\hat{n}(t) = \sum_{i\in\mathbb{L}} l_i k_i(t) \tag{6.3}$$

where $\hat{n}$ represents the accumulation.

Similarly, directly measuring the trip completion rate is often challenging, making it necessary to estimate it using readily available traffic variables that can be easily obtained. A notable finding by Geroliminis and Daganzo (2008) is that the trip completion rate exhibits a linear relationship with the trip production, with the average vehicular trip length serving as a scaling factor. Thus, we can estimate it as

$$D(t) = P(t)/\mathcal{L} \tag{6.4}$$

where $D(t)$ denotes the trip completion rate, $P(t)$ denotes the network trip production, $\mathcal{L}$ denotes the average trip length. To estimate $P(t)$, we can utilize the product of the network length and the weighted space-mean flow, denoted as $\hat{P}(t) = Lq(t)$. Therefore, we can estimate the trip completion rate as below.

$$\hat{D}(t) = \frac{q(t)}{\gamma} \tag{6.5}$$

where $\gamma = \mathcal{L}/L$ denotes a scaling factor. This relationship will be employed in Chapter 6 to estimate the trip completion rate under various disruption situations.

Table 6.1: Notation in MFD-based resilience indicators development.

Notation	Description
$\Delta^d(t)$	trip completion rate reduction at time t due to congestion
$\Delta^s(t)$	trip completion rate reduction at time t due to supply-side disruptions
R^d	resilience loss due to congestion
R^s	resilience loss due to supply-side disruptions
$D(t)$	trip completion rate at time t
D_c	optimal trip completion rate
$k(t)$	weighted space-mean density at time t
k_c	critical weighted space-mean density
$q(t)$	weighted space-mean flow at time t
q_c	optimal weighted space-mean flow
$H(\cdot)$	Heaviside step function
$P(t)$	trip production of the entire network at time t
$\mathbb{L}$	set of links in the directed network G
$\mathcal{L}$ [km]	average trip length
L	total length of all links of the network
γ	scaling factor between trip completion rate and weighted space-mean flow

6.2.2 Traffic resilience to congestion

As per the definition of transportation system resilience proposed by FHWA (FHWA, 2015), we define traffic resilience as below.

Definition 5 (Traffic resilience) *Traffic resilience represents the ability of an urban road transportation system to prepare for different kinds of disruptions, effectively serve vehicles, and recover rapidly to its optimal serving rate (i.e., trip completion rate).*

Hyper-congestion, here refers to situations of congestion resulting from extremely high demand. To make it general, traffic resilience to congestion is discussed in this study. During such disruptions, the transportation network is unable to serve users (vehicles) as efficiently as under normal operations due to the intricate interactions among vehicles in congested states and the propagation of traffic congestion. In essence, the resilience loss under congestion stems from the fact that it prevents the transportation network from operating at its optimal performance level, which is indicated by the states in the MFD where vehicle accumulation exceeds its critical value.

Considering that the primary function of transportation networks is to facilitate the mobility of people and goods (Jana et al., 2023), the trip completion rate (measured in vehicles per hour) represents the network's service rate. Hence, the performance loss under congestion can be evaluated by quantifying the reduction in trip completions. Specifically, for a given time t, the reduction in trip completions can be calculated by

$$\Delta^d(t) = (D_c - D(t))\, H(k(t) - k_c) \tag{6.6}$$

where $\Delta^d(t)$ represents the trip completion reduction, D_c is the optimal trip completion rate identified using MFD. The superscript d in $\Delta^d(t)$ refers to demand-side disruptions (here, congestion), which is used to distinguish from supply-side disruptions introduced in Section 6.2.3. $k(t)$ is the weighted average density at time t given by Equation (6.1), and k_c is the critical density. $H(k(t) - k_c)$ is an indicator function (here, a Heaviside step function) defined as

$$H(k(t) - k_c) := \begin{cases} 1 & k(t) - k_c \geq 0 \\ 0 & k(t) - k_c < 0 \end{cases} \tag{6.7}$$

This function effectively eliminates the consideration of capacity spare in the evaluation, which is align with the discussion in Section 2.3.2 regarding the shortcoming of the indicator proposed by Gao et al. (2022).

Then, the resilience loss can be quantified by the integral of trip completion reduction along the disruption period.

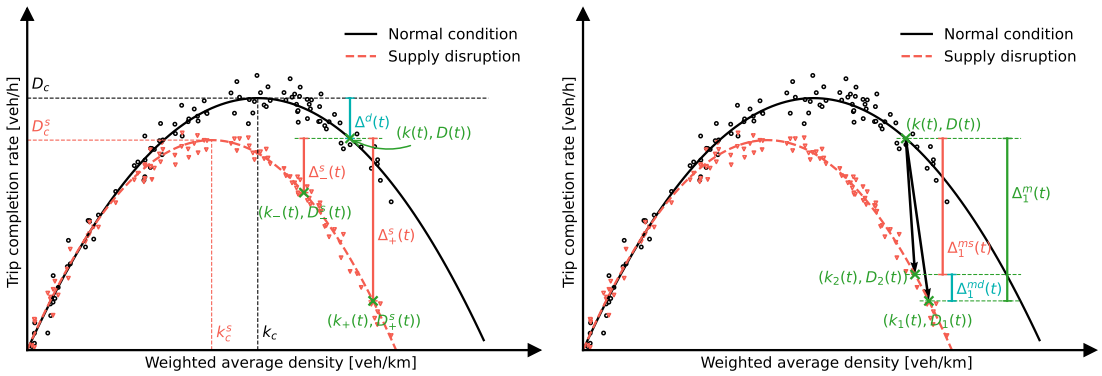
$$R^d = -\int_{t_0^d}^{t^d} \Delta^d(t)\ \mathrm{d}t = -\int_{t_0^d}^{t^d} (D_c - D(t))\, H(k(t) - k_c)\ \mathrm{d}t \tag{6.8}$$

where the time from t_0^d to t^d indicates the congestion period or the time period of interest.

Hence, given $\hat{D}_c = q_c/\gamma$, substituting Equation (6.5) into Equation (6.8), the resilience loss can be estimated by

$$\hat{R}^d = -\int_{t_0^d}^{t^d} \left(\frac{q_c}{\gamma} - \frac{q(t)}{\gamma}\right) H(k(t) - k_c)\ \mathrm{d}t \tag{6.9}$$

By utilizing Equation (6.9) with a specified γ value, we can compute the traffic resilience loss of the network contributed by congestion solely using loop detector data. The parameter γ can be estimated by utilizing floating car data or simulation data generated from the corresponding calibrated simulation model.



(a) Single type of disruption **(b)** Mixed disruption

Figure 6.2: Graphical illustration of the definition of traffic resilience to disruptions.

6.2.3 Traffic resilience to supply-side disruptions

Supply-side disruptions refer to modifications to the network structure, such as temporary closures of links, intersections, or regions due to specific events (e.g., marathons, concerts), infrastructure malfunctions caused by natural or man-made disasters (e.g., earthquakes, cyberattacks).

Unlike demand-side disruptions, supply-side ones result in resilience loss by affecting the shape of the MFD. Generally, a "shrinkage" of the MFD is anticipated. Figure 6.2a illustrates an instance where the MFD under supply-side disruptions is situated below the MFD under normal operational conditions. However, it is crucial to note that this example may not accurately represent all transportation networks. In some cities, for instance, the condition $k_c^s < k_c$ may not hold, as we will demonstrate in Section 6.3.5. Similarly, the critical accumulation may not always be smaller under supply-side disruptions.

The resilience of a transportation system under supply disruptions can be measured by comparing the dynamics of the MFD before and after the disruption. Note, the resilience loss caused by congestion under normal conditions should be removed from the calculation of the resilience loss induced by supply-side disruptions. The difference between the distances of the optimal trip completion rate to the completion rate under normal conditions and to the value under disruption conditions is employed to quantify the resilience loss solely attributable to supply disruptions. Mathematically, the trip completion reduction caused by supply disruptions, denoted as Δ^s, is calculated as

follows:

$$\Delta^s(t) = \max\left\{\left(D_c - D^s(t)\right) - \left(D_c - D(t)\right), 0\right\} = \max\left\{D(t) - D^s(t), 0\right\} \tag{6.10}$$

While $(D_c - D^s(t))$ computes the total reduced trip completion rate under the supply-side disruption, $(D_c - D(t))$ computes the reduced trip completion rate due to traffic congestion if there is no supply-side disruption. Then, their difference gives the reduced trip completion rate solely caused by the supply disruption.

The density under disruptions is usually not equal to what it should be under normal situations. For instance, as depicted in Figure 6.2a, assume the traffic state without disruptions is represented by $(k(t), D(t))$ at time t. Under disruptions, it might change to $\left(k_-(t), D^s_-(t)\right)$ or $\left(k_+(t), D^s_+(t)\right)$, indicating that the density may be lower or higher. Nonetheless, it is intuitive to expect that traffic will become denser (the latter case) if the demand is not changed as the number of available links within the network decreases. However, in practice, some travelers might cancel their trips, leading to the former case.

Then, the resilience loss due to supply-side disruptions can be quantified by

$$R^s = -\int_{t_0^s}^{t^s} \Delta^s(t)\,\mathrm{d}\,t = -\int_{t_0^s}^{t^s} \max\left\{D(t) - D^s(t), 0\right\}\ \mathrm{d}t \tag{6.11}$$

where the time from t_0^s to t^s indicates the disruption period or the time period of interest.

Similarly, given the linear relation between the trip completion rate and the weighted space-mean flow (i.e., $\hat{D}(t) = Lq(t)/\mathcal{L} = q(t)/\gamma$), we can estimate the resilience loss caused by supply-side disruptions by

$$\hat{R}^s = -\int_{t_0^s}^{t^s} \max\left\{\frac{q^s(t)}{\gamma^s} - \frac{q(t)}{\gamma}, 0\right\}\ \mathrm{d}t \tag{6.12}$$

Where $\gamma^s = \mathcal{L}^s/L^s$. We distinguish γ and γ^s here because the average trip length and the length of the available network are different under normal conditions and under supply-side disruptions due to rerouting and link closures.

6.2.4 Traffic resilience to mixed disruptions

While we have discussed the situations of demand reduction and demand unchanged under supply-side disruptions, it is often the case that sudden supply disruptions trigger demand surges. For example, emergency demand or evacuations can lead to a sudden demand increase (Safitri and Chikaraishi, 2022). This leads to a mixed disruption scenario combining supply disruptions with increased demand, which would result in more severe congestion and further impact traffic resilience.

Figure 6.2b illustrates a possible situation of a mixed disruption scenario. Assume the traffic state under normal conditions is at $(k(t), D(t))$. With the occurrence of a supply disruption, the traffic state changes to $(k_1(t), D_1(t))$. In this case, the change is attributed to both the supply disruption and the emergency demand increase. Therefore, the traffic resilience loss calculated using Equation (6.12) reflects the combined impact

of both factors, denoted as $\Delta_1^m(t)$. Assuming a disruptive traffic state without the surge in demand is represented by $(k_2(t), D_2(t))$, the traffic resilience loss solely due to the supply disruption $(\Delta_1^{ms}(t))$ is smaller than $\Delta_1^m(t)$. The difference between them is the portion attributed to congestion induced by the emergency demand, denoted as $\Delta_1^{md}(t)$.

In practice, leveraging Equation (6.12), we can obtain $\Delta_1^m(t)$ and $\Delta_1^{ms}(t)$ by running simulations with augmented demand and normal demand, respectively. Then, the loss due to demand increase can be estimated as $\Delta_1^{md}(t) = \Delta_1^m(t) - \Delta_1^{ms}(t)$. Noteworthy is that, in addition to the difference in demand, the scaling factors (γ^s) for $\Delta_1^m(t)$ and $\Delta_1^{ms}(t)$ may also differ due to vehicle rerouting and detouring.

6.2.5 Physical interpretation of traffic resilience

As per the proposed definition of traffic resilience (Definition 5) and corresponding formulations, physically, traffic resilience loss due to congestion R^d [veh] and supply-side disruptions R^s [veh] represent the cumulative number of vehicles that should have finished their trips within the respective planned time intervals under optimal operational conditions and in the absence of any supply disruptions, respectively.

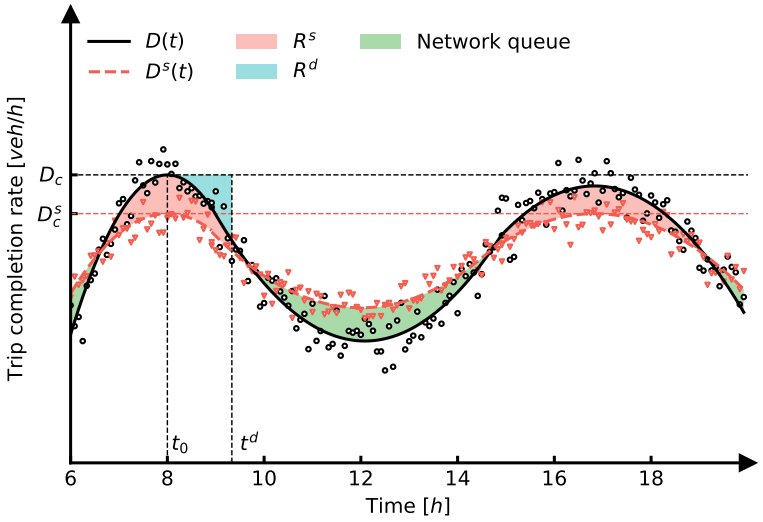


Figure 6.3: Graphical illustration of the calculation of traffic resilience.

Figure 6.3 provides an illustrative example showcasing the change of trip completion rates under normal conditions and supply-side disruptions from 6 am to 8 pm. It visualizes the corresponding traffic resilience losses. In this example, two demand peaks are observed at 8 am (morning peak) and 5 pm (evening peak). Within the congestion period from t_0^d to t^d (assuming that the system returns to be uncongested at t^d), the traffic resilience loss due to congestion is measured by the area between D_c and the curve of $D(t)$. Conversely, the traffic resilience loss due to supply-side disruptions is measured by the area between $D(t)$ and $D^s(t)$ when $D(t) > D^s(t)$. When the demand reduces to a certain level (below the service rate under supply-side disruptions), $D^s(t)$ will exceed

$D(t)$. In this case, the system is serving the "cumulative vehicles" that were delayed by the disruption. We refer to these delayed vehicles as the network queue, as indicated in Figure 6.3. In other words, the network queue ([veh]) at time t is the remaining portion of traffic resilience loss ([veh]) that has not been served.

6.2.6 Normalization and discretization

Comparing the resilience of transportation networks across different cities has been a longstanding interest within the transportation community (e.g., Ganin et al., 2019; Kurth et al., 2020; Yin et al., 2023). However, such comparisons require normalization due to the inherent differences in city scales, including factors such as area, population, and car ownership. To address this, we utilize the optimal trip completion rate, which defines the optimal functionality of the system, as a normalization factor for the trip completion reductions. This allows us to calculate the normalized resilience losses, which are given by

$$\tilde{R}^d = -\int_{t_0^d}^{t^d} \left(\frac{D_c}{D_c} - \frac{D(t)}{D_c}\right) H(k(t) - k_c)\ \mathrm{d}t = -\int_{t_0^d}^{t^d} \left(1 - \frac{q(t)}{q_c}\right) H(k(t) - k_c)\ \mathrm{d}t \quad (6.13)$$

$$\tilde{R}^s = -\int_{t_0^s}^{t^s} \max\left\{\frac{D(t)}{D_c} - \frac{D^s(t)}{D_c}, 0\right\} \mathrm{d}\,t = -\int_{t_0^s}^{t^s} \max\left\{\frac{q(t)}{q_c} - \frac{\gamma q^s(t)}{\gamma^s q_c}, 0\right\}\ \mathrm{d}t \quad (6.14)$$

The traffic resilience formulations discussed above are derived for continuous observations. However, in practice, traffic dynamics are collected and aggregated into discrete time intervals. In this case, we can calculate the trip completion reduction within each time interval and then estimate the total traffic resilience losses by summing up the reductions over all intervals. The estimation of total traffic resilience losses in discrete time intervals can be expressed as follows:

$$\hat{R}^e = -\frac{T}{2}\sum_{t=1}^{N_t}\left(\hat{\Delta}^e(t) + \hat{\Delta}^e(t-1)\right) \quad (6.15)$$

where $e \in [d, s]$ indicates the type of disruption events, T is the length of time intervals (assuming all time intervals share the same length), and N_t is the total number of time intervals within the period of interest. We assume that the period of interest starts before the disruption events, such that we have $\hat{\Delta}^e(0) = 0$.

6.3 Case study and results

To examine the proposed traffic resilience indicators, we implement case studies in two cities of different types of road network topology/layout. Specifically, we selected the city center of Munich, Germany, which features a central ring network, and the city center of Kyoto, Japan, known for its typical grid network. See detailed description of these two networks in Section 3.4.2.3 and Section 3.4.2.2. Residential links are excluded from both networks.

6.3.1 Simulation setup

Simulation of Urban MObility (SUMO) (Lopez et al., 2018), an open-source traffic simulator, is used to generate traffic dynamics for estimating the MFDs under various scenarios that are expounded on in Section 6.3.3. All scenarios are simulated at the mesoscopic resolution with a non-iterative dynamic stochastic user assignment method that approximates the Dynamic User Equilibrium (DUE) (Lu et al., 2023b). To reduce the influence of the stochasticity in simulations, results from 10 replications are used to analyze the MFD dynamics of each scenario and estimate relevant variables, such as the critical average density and the optimal weighted space-mean flow. For each scenario, the traffic between 6 am and 10 am are simulated, with the first and last half an hour as the simulation warm-up and dissipation periods, respectively.

6.3.2 Mesoscopic simulation model calibration

For ensuring reliable simulation results, we also conduct a calibration on the mesoscopic simulation models. For the calibration of the mesoscopic simulation model, we mainly focus on the OD matrices calibration by using the traffic counts collected by loop detectors. Normalized root mean square error (RMSN) is used to evaluate the performance of calibration, which is given by

$$\mathrm{RMSN} = \frac{\sqrt{N \sum_{i=1}^{N} (\hat{y}_i - y_i)^2}}{\sum_{i=1}^{N} y_i} \tag{6.16}$$

where y_i and $\hat{y}_i$ are the observed and simulated traffic counts at detecting location i, N denotes the total number of detecting locations.

For the Kyoto model, we use the movement matrices of *Docomo* (a Japanese mobile phone operator) users as the initial guess of the vehicle mobility patterns. Considering the difference in the scales of the two groups, we first find out the scaling factors resulting in the smallest RMSN for each time interval by simply trying different convex combinations of predefined lower and upper bounds. This step improves the matrices from an RMSN error of 1.2 to 0.74. It indicates that the mobility patterns of mobile phone users can resemble that of vehicles to a certain extent. We note that the study area is only a part of the city of Kyoto. As a result, the amount of through traffic is very different from the ones extracted from mobile phone data (Lu et al., 2023b) due to : (i) All vehicles can only use the network provided to finish their trips which is inconsistent with reality in which some can reach their destinations by the paths outside this network; (ii) Vehicles from the area outside the study area have not been counted. To address this issue, we first apply the simultaneous perturbation stochastic approximation (SPSA) algorithm (Spall, 1998) to correct the OD demand from/to the outermost zones to additionally measure the demand from/to external zones. Utilizing the corrected OD matrix as the prior, we then employ the PC-SPSA algorithm (Qurashi et al., 2022) to calibrate the whole OD matrix. Finally, the average RMSN reduces to a satisfactory level of 0.41.

For the Munich model, the reader is referred to the authors' previous studies in Lu et al. (2021) and Dadashzadeh et al. (2021) for more details about its calibration.

6.3.3 Demand variations and supply-side disruption scenarios

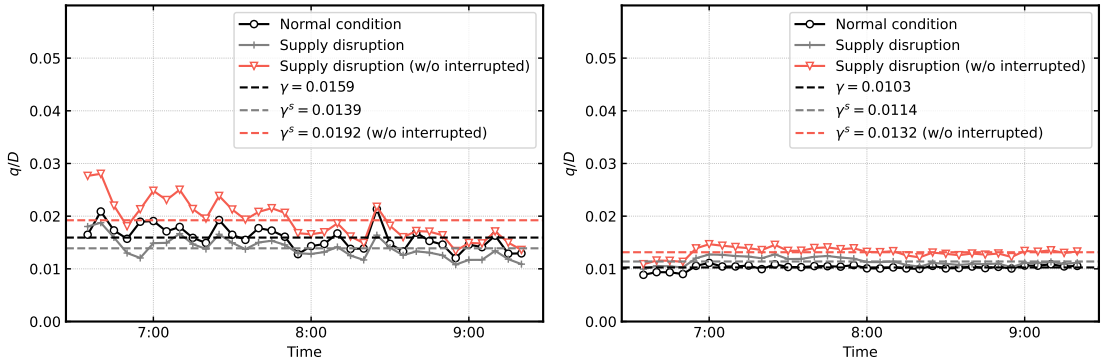
To demonstrate the MFD dynamics under different demand scenarios and to validate the underlying assumptions in the proposed approach, we consider three different travel demand levels: small demand (SD), medium demand (MD), and large demand (LD). The SD scenario represents the original demand level, while the MD and LD scenarios correspond to 1.2 and 1.5 times of the original demand, respectively. For each study area, all three demand scenarios are simulated. However, it is worth mentioning that, in real applications, the disrupted demand should be estimated correspondingly with available data, either conventional or emerging data (Mahajan, 2023), for instance, by leveraging the approaches presented in Chapter 4 and Chapter 5.

The design of supply-side disruptions (SSD) is explained in Section 3.4.2.2 and Section 3.9a. The SSD scenarios of the two cities result in comparable proportions, accounting for 7% and 8% of closed links in their respective networks. Also, the closed links correspond to the busiest areas within the respective networks. This allows for a direct comparison of traffic resilience to supply disruptions between the two networks.

6.3.4 Trip completion rate estimation using a proxy

Considering the changes in the average trip length and the total length of the network under supply-side disruptions, the linear scaling factor is very likely to be different. In this section, we used the SSD scenarios described in Section 6.3.3 as an example to examine this estimation approach. To estimate the appropriate scaling factor γ^s for each city's SSD, we conduct simulations with the closed links across the entire simulation period. This allows us to generate the necessary information for the γ^s. Figure 6.4 demonstrates the average ratio of $q(t)$ to $D(t)$ from 10 simulation replications in different time intervals, excluding the warm-up and dissipation periods. Noteworthy, apart from the canceled trips originating from or heading to the closed area, there are also some trips cannot be carried on due to the absence of alternative routes, which are called interrupted trips here. Figure 6.4 presents two estimated γ^s values for each city: one considering all trips (in grey) and another one excluding interrupted trips (in red). For Munich, the former is smaller than γ while the latter is greater. In Kyoto, both are greater than γ. In general, γ^s should be greater than γ due to the network length decline and the trip length increase (because of detouring). However, this is not necessary always true. If a large amount of long-distance trips were canceled for no available route alternatives, it is also possible to observe a decrease in the average trip length. In this case, the decrease proportions of the average trip length and network length will determine whether γ^s is greater than γ. The smaller γ^s obtained in Munich indicates that the average trip length under SSD reduces with a proportion of reduction smaller than that of the total network length.

In order to explain the factors contributing to the difference, we also plot relevant trip statistics in Figure 6.5. As can be seen, approximately 12% of trips are interrupted in Kyoto, while this number increases to 21% in Munich. This disparity can be attributed to the grid structure of the Kyoto network, which provides more route alternatives between



(a) Gammas of Munich **(b)** Gammas of Kyoto

Figure 6.4: Estimation of γ under normal conditions and the supply disruption.

every OD pair, enabling more trips to be rerouted after network disruption compared to the Munich network. This is also consistent with the finding in Zhang et al. (2015). Furthermore, the grid structure can also ease the vehicle detouring, represented by the minor average trip length increase. The average trip length of *completed* trips (i.e., trips reaching their destination) is 1.05 multiples of that under normal conditions in Kyoto, while this number is 1.27 in Munich. However, if we include the interrupted trips in the calculation, both cities reach a number of about 0.95, i.e., a reduction of about 5% in the average trip length. This value serves as a reference for the comparison of γ^s in the two cities. Figure 6.5 also depicts the percentage of the length of closed lanes in the complete network. Although we designed the scenarios by closing a similar proportion of links (7% for Kyoto and 8% for Munich), the total length of closed lanes comprises 8.98% (greater than 5%) and 5.68% (comparative to 5%) of the complete network of Kyoto and Munich, respectively. This leads to the difference of the change in γ^s between the two cities. It is worth mentioning that the errors between γ's and $\mathcal{L}/L$ can be expressed by the estimation of the production using q. To eliminate the noise introduced by the interrupted trips, we improve the estimation by excluding those trips. The corrected values are also depicted in Figure 6.4 and will be utilized in the following evaluation.

6.3.5 MFD dynamics analysis

Figure 6.6 shows the MFD dynamics (aggregated every 5 min) of the scenarios described in Section 6.3.3. In the case of Munich (Figure 6.6a), in terms of demand variations, we can see that neither scenario SD nor MD exceeds the critical point, indicating the absence of network-wide congestion. We can also see that the dynamics become more unstable as traffic approaches the critical point, which is consistent with the analysis in Gao et al. (2022). Furthermore, as shown in Figure 6.7a and 6.7b where results from different replications are averaged (with standard deviations shown), a clockwise hysteresis loop is observed in both scenarios, but the size of the hysteresis loop is almost

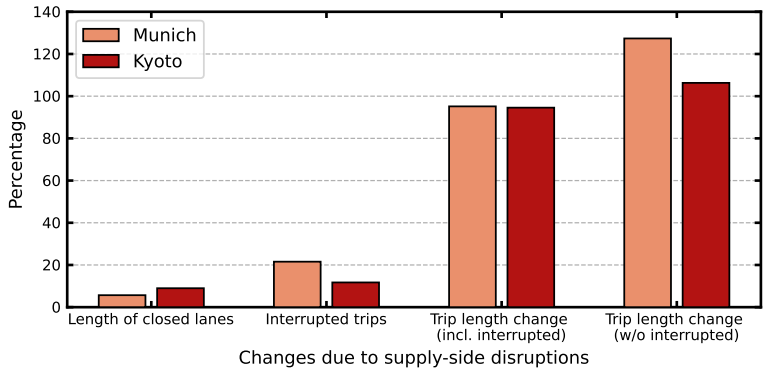
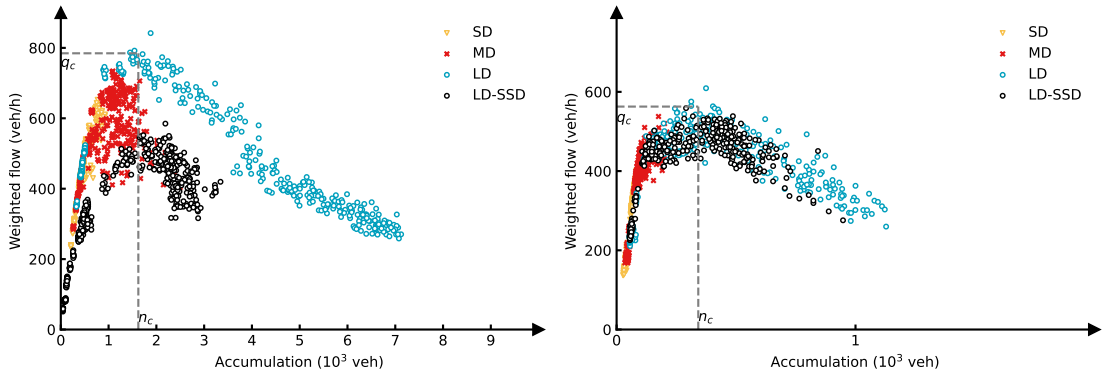


Figure 6.5: Changes in supply and demand due to supply-side disruptions.

doubled in MD compared to SD. The larger standard deviations observed in MD further confirm the presence of an unstable state around the critical point.



(a) Munich network **(b)** Kyoto network

Figure 6.6: MFD dynamics of the scenarios of investigation.

On the other hand, in scenario LD (Figure 6.7c), gridlock occurs in the network due to the over-saturated traffic. We find that the MFD dynamics are less dispersed in the vicinity of the critical point. Additionally, we plot the MFD dynamics for the scenario with supply disruption and large OD demand (LD-SSD) in Figure 6.6a. As expected and consistent with findings in Kim and Yeo (2017) and Gao et al. (2022), a noticeable reduction in weighted space-mean flow is found at the same accumulation level. However, the critical accumulation, in this case, denoted as n_c^s, is comparative to that of the normal system n_c. It indicates that using congestion deviation (i.e., $n(t) - n_c$) to model the resilience loss due to supply-side disruptions as in Gao et al. (2022) will lead to inaccurate estimation.

Figure 6.6b shows the traffic dynamics of Kyoto under different scenarios. Similar to the Munich case, Kyoto also experiences fluctuations in demand variation scenarios.

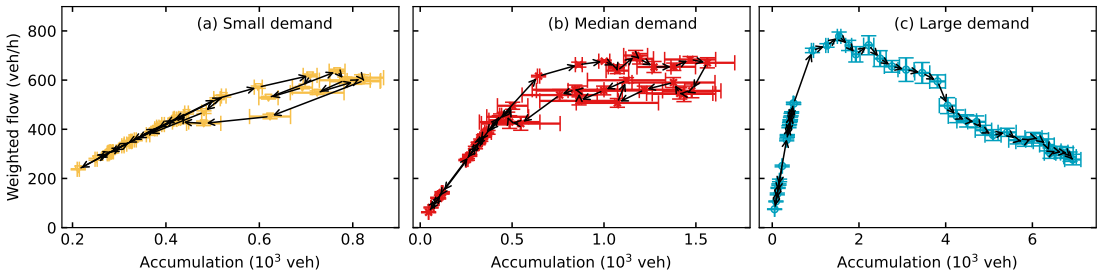


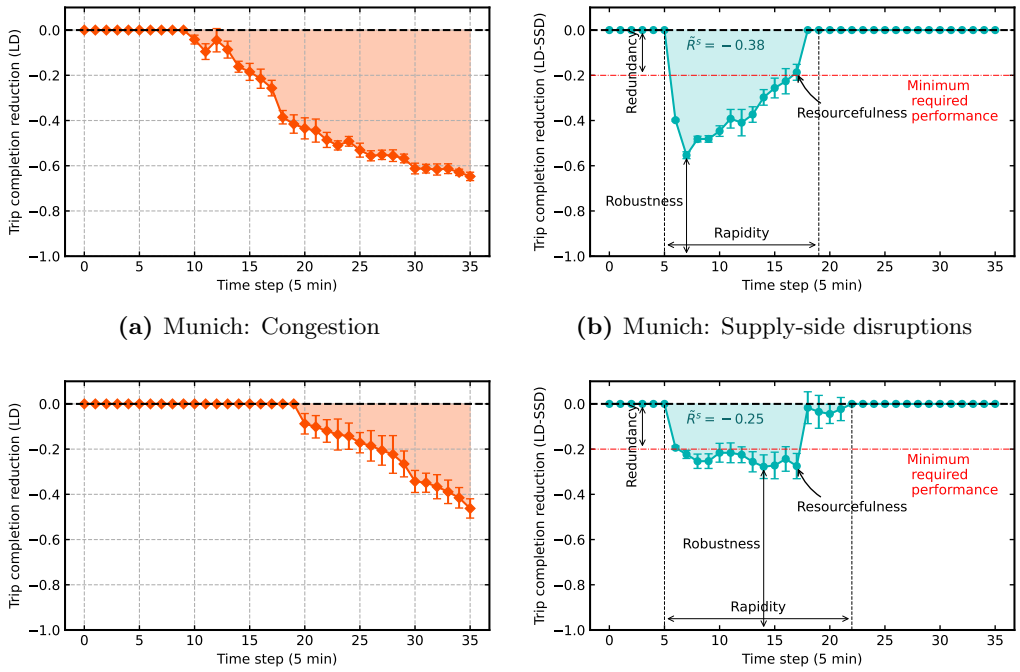
Figure 6.7: Development of MFD dynamics due to demand variations in Munich.

In contrast, in terms of the LD-SSD, Kyoto almost suffers no degradation in the MFD shape. Therefore, it is unreasonable to directly use the area between the normal MFD curve and the disruptive MFD curve as the measurement of resilience loss as in Kim and Yeo (2017), as it will result in no or a very small resilience loss in Kyoto, which is incorrect as we showed in Section 6.3.6. Nevertheless, it is essential to note that differences do exist between the points of the same time interval under different scenarios, as explained in Section 6.2. This finding reaffirms the notion that the influencing mechanisms of congestion and supply-side disruptions are distinct, and therefore the traffic resilience to them should be discussed separately. It is thus reasonable to state that the resilience indicators proposed in this study are more reliable and robust for evaluating different types of disruptions and transportation systems.

6.3.6 Traffic resilience evaluation

In this section, two scenarios are considered for testing the system performance: Scenario LD is employed to assess the system's behavior under extreme demand situations, such as emergency travel demand, while scenario LD-SSD is utilized to test its performance in the presence of network or infrastructure malfunctions. For comparison purposes, we adopt the formulation of normalized traffic resilience (i.e., Equations (6.13) and (6.14)) for the evaluation. The estimated γ's presented in Section 6.3.4 are used for the normalization. The resilience loss curves (i.e., the normalized trip completion rate reduction over different time intervals) for congestion and supply-side disruptions in Munich are given in Figure 6.8a and 6.8b, respectively. The corresponding curves for Kyoto are given in Figure 6.8c and 6.8d.

We note that the complete life cycle of the influence of congestion is not present due to the limitation of simulation duration (due to lack of demand data). The recovery process is missing in the LD experiments, which prevents us from drawing definitive conclusions regarding which city is more resilient to demand variations, especially when observing similar reduction speeds during the disruption phase in both cities. On the contrary, from the curves for supply-side disruptions, it is evident that Kyoto is more resilient to supply-side disruptions (here, partial link closure), represented by a smaller area enclosed by the "resilience triangle". The normalized traffic resilience losses, $\tilde{R}^s$, are -0.38 and -0.25 for Munich and Kyoto, respectively.



(a) Munich: Congestion

(b) Munich: Supply-side disruptions

(c) Kyoto: Congestion

(d) Kyoto: Supply-side disruptions

Figure 6.8: Traffic resilience under congestion and supply disruptions (large demand scenario).

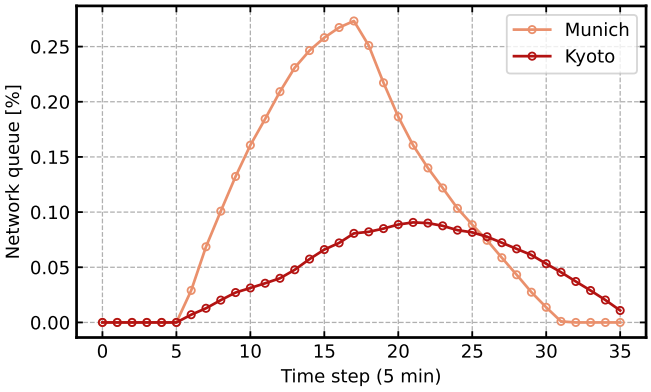


Figure 6.9: Network queue due to the supply disruption (LD-SSD).

As per the "4R" resilience properties (i.e., robustness, redundancy, resourcefulness, and rapidity) proposed in Bruneau et al. (2003) and characterized in Wan et al. (2018), we can draw the following conclusions: (i) A greater minimum value obtained in Kyoto represents stronger robustness in dealing with disruptions. (ii) A longer during-disruption duration (including response and recovery phases) in Kyoto implies that Munich outperforms Kyoto in terms of the property of rapidity. (iii) The resourcefulness property affects the shape of the system functionality curve during disruption. These two cities differ significantly from each other in this regard. To be specific, Munich gradually recovers after reaching the worst performance point, whereas Kyoto oscillates back and forth around that point until links are reopened. However, it is worth noting that the worst point in Kyoto is much better than in Munich, indicating that even during the disruption period, a stable "new equilibrium" can easily form in Kyoto. Nonetheless, future works on developing a specific quantitative measurement for the resourcefulness property would still be desirable. (iv) Due to the lack of a specific definition of the *minimum required performance* here, redundancy cannot be specified. If a minimum required performance is defined as the example we provide in Figure 6.8b and 6.8d, i.e., -0.2 for each, the redundancy can be measured by its absolute value. However, the redundancy property can also be evaluated by the existence of optional routes for OD pairs. In this regard, grid-like networks usually outperform other kinds of networks. It is essential to consider these various resilience properties while evaluating the performance of transportation networks under different disruptions, as they provide valuable insight into how cities respond to and recover from disturbances.

Given that the network queue can be regarded as a part of traffic resilience loss under supply disruptions, the trends observed in the network queue curves presented in Figure 6.9 mirror those of the functionality curves illustrated in Figure 6.8. For comparison purposes, the network queues have also been normalized by the respective critical traffic flows q_c. While Munich exhibits a significantly longer network queue compared to Kyoto, it is noteworthy that the Kyoto system requires more time to clear the network queue.

The analysis presented above clearly highlights the crucial role of network topology for traffic resilience. However, it is important to acknowledge that traffic resilience is not solely determined by network topological attributes; it is also significantly influenced by traffic dynamics, as evidenced by the earlier analysis. Therefore, the question of how much network topological attributes can account for traffic resilience necessitates further investigation and exploration.

6.4 Conclusions and future work

This chapter investigates the problem of evaluating traffic resilience of urban road transportation systems. Traditional approaches have relied on static topology-based indicators, such as accessibility, or simple aggregation of trip information, such as average speed. However, these indicators have limitations: the former fails to capture the dynamic nature of traffic, while the latter is sensitive to travel demand. Moreover, neither

approach adheres to the definition of system resilience as they cannot be considered as a inherent property of networks. Recently, several indicators based on the concept of MFDs have been developed to measure the resilience of urban transportation systems from multiple perspectives. Unlike other resilience indicators, MFD-based indicators possess the advantage of being network properties, as MFD represents an intrinsic characteristic of a homogeneously congested network. Nonetheless, existing MFD-based indicators have certain drawbacks, such as insufficient consideration of network structure changes, neglecting the detouring or rerouting of vehicles, or relying on inaccurate reference points.

To overcome the aforementioned limitations, this chapter introduces a novel approach that offers a comprehensive evaluation of traffic resilience. Our approach specifically addresses the distinct influencing mechanisms of congestion and supply-side disruptions. Notably, while supply-side disruptions may alter the shape of the MFD, traffic congestion does not have the same effect. We separately discussed traffic resilience to congestion and supply-side disruptions and built the respective indicators accordingly. Additionally, we outline a methodology for discretizing and normalizing the calculation of traffic resilience loss, enabling practical applications and facilitating comparative analyses. By using the proposed indicators, traffic resilience loss can be physically interpreted as the cumulative number of vehicles that should have finished their trips within the respective time intervals if the transportation system/network were operating optimally under normal conditions (without infrastructure malfunction).

We conducted experiments in two cities — Munich, Germany, and Kyoto, Japan, which possess different network topologies. The experiment results revealed that Kyoto's grid-like network demonstrates greater resilience to supply-side disruptions compared to Munich's central ring structure. Several factors contribute to this disparity: (i) Grid-like networks offer more route alternatives between every OD pair so that more trips can be rerouted after network disruption; (ii) The grid structure can also ease the vehicle detouring.

We acknowledge the potential occurrence of mixed scenarios resulting from the interplay between a supply-side disruption and emergency demand (attributed to recovery actions or public panic), as discussed in Section 6.2.4. However, the analysis of related scenarios in the experiment results is limited. Future studies can be undertaken to develop a reliable approach to design plausible mixed disruption scenarios and explore the interaction and interdependence between traffic resilience of these two types of disruptions in such scenarios.

Chapter 7

Regression analysis between traffic resilience and network topology

The content of this chapter has been partially presented in the following works:

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2023b). Simulation-based policy analysis: The case of urban speed limits. *Transportation Research Part A: Policy and Practice*, 175:103754.

Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2024c). Traffic resilience quantification based on macroscopic fundamental diagrams and analysis using topological attributes. *Reliability Engineering & System Safety*, 247:110095.

7.1 Introduction and research contributions

Topological attributes play a critical role in the resilience property of transportation networks (Zhang et al., 2015; Pan et al., 2021; Hao et al., 2023). For instance, attributes such as the Alpha index and betweenness centrality can effectively characterize networks' connectivity and accessibility, respectively. It follows that they have been extensively employed to define and assess transportation system resilience. However, these metrics fail to capture the traffic dynamics of the system, and applying a single metric of them for the assessment may lead to unreliable conclusions. To address this limitation, some studies (such as Levinson, 2012; Parthasarathi, 2014; Huang and Levinson, 2015; Kurth et al., 2020; Wang et al., 2023; Yin et al., 2023) have been proposed to model the relationship between topological attributes and traffic dynamics so as to understand to which extent network topology can determine the operation of the associated transportation system. However, all of them have at least one of the following limitations: (i) Different transportation networks are included in a single model regardless of the topology type/style (e.g., grid-like, scale-free); (ii) The influence of travel demand on the dependent variable is overlooked.

In this chapter, we are interested in the relationship between topological attributes and traffic resilience to supply-side disruptions. The reason why congestion is not included is that it often refers to recurring events, while supply-side disruptions pertain to non-recurring events. Also, only supply disruption events will change the network topology. To model the relationship, observations recording the network structure and traffic states under different disruptive events are required. However, given the rarity

of such events, it is nontrivial to collect sufficient data. Fortunately, traffic simulation provides an economic and reliable way to address this problem. Therefore, this chapter aims to develop a simulation-based synthetic scenario generation procedure to simulate the influence of various supply-side disruptions. From these synthetic disruption scenarios, we can make one-to-one correspondence between the metric values of topological attributes and traffic resilience losses. Specifically, in this chapter, we conduct a regression analysis on their relationship with the traffic resilience loss evaluated using the indicators developed in Chapter 6.

This chapter offers the following contributions:

- In this chapter, we build a regression model to describe the relationship between topological attributes of transportation networks and traffic resilience. With this regression model, traffic resilience can be roughly predicted based on the network structure, eliminating the necessity for traffic flow information.
- We construct a simulation-based procedure for systematically creating synthetic supply disruption scenarios.
- We compare different MFD-based indicators, which evidences the advantages and comprehensiveness of our traffic resilience indicators.

The rest of the chapter is structured as follows. Section 7.2 presents the synthetic scenario generation procedure, and the regression model. Section 7.3 compares different MFD-based indicators and analyzes the regression result. Finally, Section 7.4 concludes the chapter and suggests future research directions.

7.2 Methodology

This section first describes how supply-side disruptions influence vehicle behaviors and develops a simulation-based synthetic scenario generation procedure for supply-side disruptions. Then, we summarize the main network topological attributes and corresponding metrics. Finally, a regulated regression model is presented to estimate the relationship between these attributes and traffic resilience.

7.2.1 Impact of supply-side disruptions on vehicle behaviors

Vehicles typically choose routes based on their utilities, which is evaluated by travel costs. Travel time, as the most commonly used metric of travel cost, can be employed to gauge changes in route efficiency following the occurrence of supply-disruption scenarios. However, comprehending the impact of supply-side disruptions on route travel time is not straightforward, as it depends on an equilibrium formed by several conflicting effects. These conflicts can unilaterally determine the ultimate performance and thus require explicit consideration. Among other factors, modeling and estimating these conflicts also constitute one rationale for utilizing traffic simulations in synthetic scenario generation.

The logic of the influence of supply-side disruptions on vehicle behaviors is illustrated in Figure 7.1. Generally, supply-side disruptions are modeled as link closures or speed

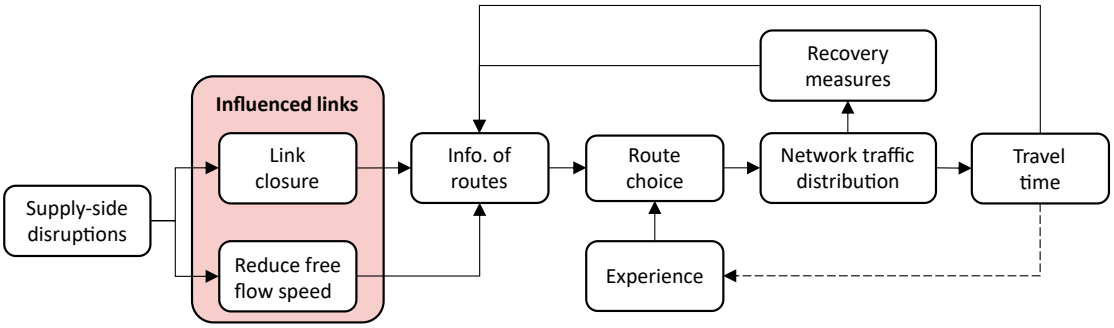


Figure 7.1: Influence of supply-side disruptions on vehicle behaviors.

reductions in simulations within the literature. Link closures result in route breaks, while speed reductions alter traffic states on affected links. These state changes impact vehicle route choices, subsequently altering traffic distribution across the network. Traffic assignment, in turn, affects the overall network traffic efficiency, reflected in disparities in travel time before and after the disruption. To improve the traffic states, recovery measures are needed. This topic is omitted here and will be discussed in Chapter 8 and Chapter 9. It is important to note that the network requires time to reach a new steady state. The route choice behavior of a vehicle is dependent on the latest information on routes, such as travel time, recovery measures, and past experiences. In other words, the traffic distribution across the network continues to change until a new equilibrium between network supply and route choice is reached.

Conventionally, the iterative simulation-based traffic assignment is used to compute the user equilibrium state. However, in reality, a proportion of people (e.g., non-routine vehicles) conform to the assumption of dynamic stochastic assignment regarding route choice behavior, continuously using navigation to achieve dynamic user optimum under a given stochastic state of the network. As a result, simulations should be conducted with imperfect DUE assignment, or more specifically, with a combination of DUE and dynamic stochastic user assignment.

7.2.2 Synthetic scenario generation procedure

In this section, we propose a systemic method to create synthetic supply disruptions, by leveraging advanced traffic simulators, for estimating the relationship model, which can also address the limitations in previous studies. We introduce the notation in Table 7.1 for the explanation in the following text of this chapter.

Figure 7.2 demonstrates the flowchart of the procedure for synthetic supply disruption scenario generation. For simplicity, we consider supply disruptions with random link closures. However, it is easy to extend the procedure to other types of disruptions. A batch of scenarios should be generated for a specific disruption configuration to ensure statistical significance. The following description explains the process of creating a synthetic sample dedicated to a certain city network (we denote its complete form by G) and a certain demand matrix (denoted by M).

Step (1) The process starts with a certain percentage number $p \in [0, 1)$, indicating the percentage of links that are blocked or damaged due to the disruptive event.

Step (2) With a certain random seed r, a disruption scenario $\mathbb{S}$ is created, by randomly sampling the links to be closed. Essentially, a disruption scenario $\mathbb{S}$ can be expressed as a set of closed links.

Step (3) Relevant topological attributes $\mathbf{x}$ describing the network topology are derived from the damaged network $G(\mathbb{S})$.

Step (4) Run traffic simulations ($\mathcal{S}$) using the damaged network $G(\mathbb{S})$ and the demand matrix M to generate traffic dynamics $Y(\mathbb{S})$. Note that in order to obtain statistically significant results, multiple simulation replications are necessary for each disruption scenario. Furthermore, the portion of demand originating from or terminating at the closed links is removed during simulations, representing the canceled trips in reality.

Step (5) From the traffic dynamics of all simulation replications, we can estimate the mean traffic resilience loss $R^s(\mathbb{S})$. Finally, the tuple $(\mathbf{x}, R^s)$ constitutes a synthetic observation.

By repeating this process for different r's under the same p and for different p's, we can create a synthetic dataset (X, R) of disruption observations.

Table 7.1: Notation in regression analysis between traffic resilience and topological attributes.

Notation	Description
r	random seed
$\mathbb{S}$	disruption scenario — a set of closed/damaged links
G	directed graph of the complete network
$G(\mathbb{S})$	directed graph of the disruptive network of scenario $\mathbb{S}$
$\mathbf{x}$	vector of topological attributes
$\mathcal{S}$	traffic simulator
M	origin-destination demand matrix
Y	simulation outputs
X	dataset of topological attributes
R^s	traffic resilience loss due to supply disruptions
R	vector of traffic resilience losses
θ	vector of parameters of the regression model
ρ	hyper-parameter related to the l_1-norm regularization term
$f(\mathbf{x}; \theta)$	linear model with explanatory variables $\mathbf{x}$ and parameters θ

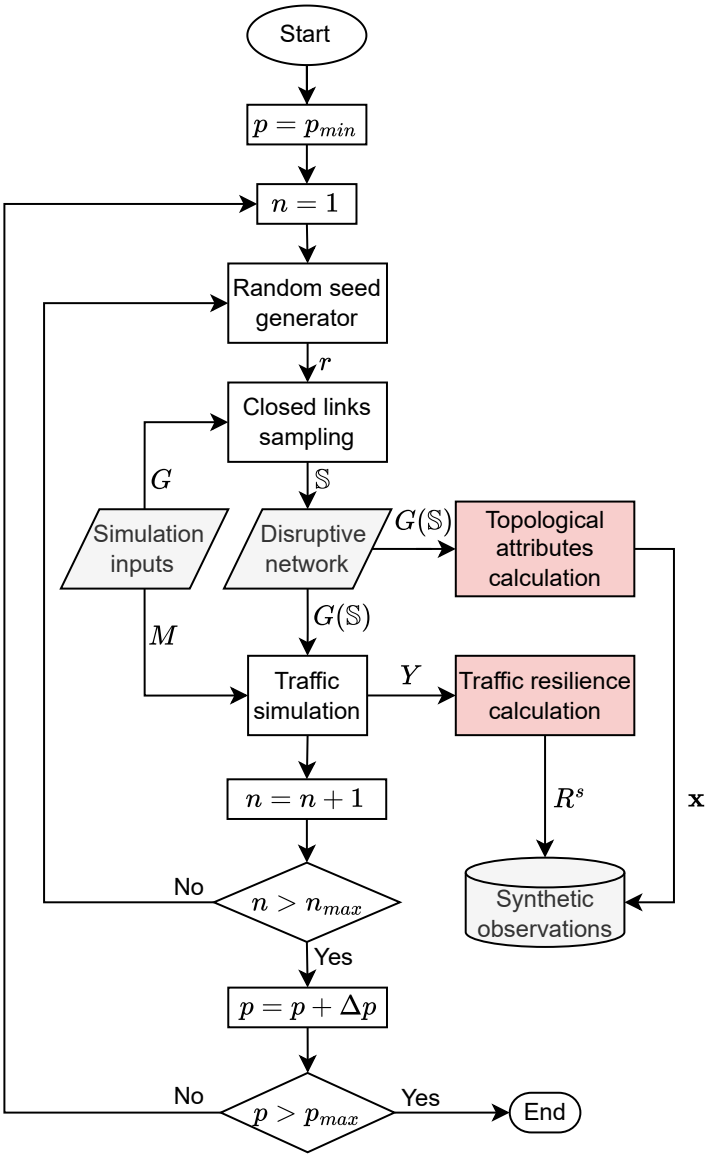


Figure 7.2: Flowchart for the synthetic scenario generation procedure.

7.2.3 Transportation network topological attributes and metrics

Table 7.2 lists some selected typical topological attributes that are commonly considered in complex network topology analysis together with the respective metrics (Hagberg et al., 2008; Zhang et al., 2015; Boeing, 2017). Note that while transportation networks are directed graphs, some metrics are defined for undirected graphs only. For calculating these metrics, the corresponding undirected graph of the network is used. The listed metrics can explain the properties of the network across multiple aspects and are used in the regression analysis in this study. Readers are referred to Boeing (2017) and Wang et al. (2023) for more details about these attributes and metrics.

7.2.4 Variable normalization and regression model development

Using the synthetic dataset, a regression model can be estimated to describe the relationship between topological attributes and traffic resilience loss for a certain type of network topology. Recall that traffic resilience loss is a relative value derived from the difference between the traffic dynamics under disruptive and normal operations. Consequently, in order to measure how the degree of resilience relates to the changes in topological attributes, the explanatory dataset X should also be normalized with the corresponding values of the complete networks and transformed into relative percentage values. Mathematically, for a variable X_i, it is transformed into

$$Z_i = \frac{X_i}{X_i^c} - 1 \tag{7.1}$$

where X_i^c is a scalar representing the i-th topological metrics computed from the complete network, Z_i represents the transformed variable of X_i. We emphasize that this transformation is necessary for obtaining better model interpretability, and this treatment also differs our model from the previous ones.

Employing a regularized linear regression, we can estimate the parameters by least squares as below:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \sum_n \left(f(\mathbf{x}_n;\theta) - R_n^s\right)^2 + \rho|\theta|_1 \tag{7.2}$$

where $\mathbf{x}_n$ denotes the n-th sample in X and R_n^s is the corresponding resilience loss, θ denotes the vector of parameters to be estimated. $f(x;\theta)$ is a linear function with respect to θ and its specific form depends on the transportation network of interest and the variable selection procedure. $|\theta|_1$ denotes the l_1-norm regularization of the coefficients. The l_1-norm term is also served as a built-in feature selection method, thereby leading to sparse solutions. ρ is a hyper-parameter, indicating the weight of the penalty term. Equation (7.2) represents a lasso model.

7.3 Case study and results

We utilize the same study areas of Munich city and Kyoto city to the ones in Chapter 6 for case studies. We also adopt the same simulation setups therein for the following experiments. In this section, we first introduce the setups of the simulation-based synthetic

Table 7.2: Typical topological attributes.

Attribute	Metric	Description
Basic statistics		
Network statistics	Number of nodes	
	Number of links	
	Average node degree	
	Total length of all edges	
	Average length of all edges	
Directed graph		
Assortativity	Degree assortativity	The assortativity coefficient is the Pearson correlation coefficient of degree between pairs of linked nodes.
	Average neighbor degree	The average neighbor degree of a node measures the average degree of all nodes within the neighborhood of a specific node.
Centrality	Average degree centrality	The average of the fraction of nodes that a node is connected to.
	Load centrality	The average of the fraction of all shortest paths passing through a node.
	Edge load centrality	The average of the fraction of all shortest paths passing through an edge.
	Harmonic centrality	The average of the sum of the reciprocal of the shortest path distances from all other nodes to a node.
Connectivity	Alpha index	Ratio of existing circuits to the maximum possible circuits.
	Beta index	Ratio between the number of links and number of nodes.
	Gamma index	Ratio of the number of links to the maximum possible number of links.
Reciprocity	Reciprocity	Ratio of the number of edges pointing in both directions to the total number of edges.
Undirected graph		
Clustering	Average clustering coefficient	The clustering of a node is the fraction of possible triangles through that node that exists.
Efficiency	Global efficiency	The efficiency of a pair of nodes in a graph is the multiplicative inverse of the shortest path distance between the nodes.

network disruption scenario generation procedure in our case studies. Then, we compare the our indicators with previous MFD-based indicators reviewed in Section 2.3.2, which lends support to the selection of our indicators for the regression analysis. Finally, we explore the relationship between topological attributes and traffic resilience indicators.

7.3.1 Synthetic network disruption scenario design

In order to understand the relationship between traffic resilience and topological attributes, we want to estimate a regression model between them with traffic resilience loss as the dependent variable. Unlike the previous studies incorporating transportation networks of different cities into a single model regardless of the type of topology, we acknowledge that their relationship may differ across topology types or cities, based on the findings from case studies in Munich and Kyoto. Therefore, we adopt the scenario generation procedure proposed in Section 7.2.4 to generate numerous synthetic disruption scenarios for each city separately, leveraging the advance in traffic simulators and their interfaces.

More specifically, for each city, we consider a range of percentages from 2% to 20%, with a 2% interval, resulting in 10 different percentage values p. For each p, 100 disruption scenarios are constructed by randomly removing p percent of links from the complete network for each scenario. Finally, 1000 scenarios are constructed. It is worth mentioning that, isolated nodes will be removed from the network as well, if any, after removing a set of links. Vehicle trips without route connections, after the link removal, are considered as *interrupted trips*. For each scenario, as described in Section 7.2.4, the topological attributes listed in Table 7.2 will be calculated using the disruptive network, and the corresponding traffic resilience loss will be calculated using the traffic dynamics generated from SUMO. Note, here traffic dynamics only include the measurements collected by loop detectors, i.e., traffic flow, occupancy, and mean speed, to mimic real-world conditions. All traffic measurements are aggregated every five minutes. Furthermore, the scaling factor γ^s relating to the trip completion rate and weighted flow have to be distinguished among scenarios to ensure accurate comparisons. All scenarios are run for the time period between 6 am and 10 am with the calibrated demand as in Chapter 6.

7.3.2 Comparison of MFD-based traffic resilience indicators

In Section 2.3.2, we analyzed the limitations of previous MFD-based resilience indicators and highlighted the advantages of our proposed indicators. To validate our analysis, we compare the curves of these normalized resilience indices under the LD-SSD scenario of Munich in Figure 7.3. We omit the indicator proposed by Kim and Yeo (2017) as it is static and cannot provide a comparative analysis. This also indicates that their establishment does not stem from the system functionality curve used in the system resilience definition.

Upon comparison, we observe that our index and Amini's index (Amini et al., 2018) show a very close pattern during the disruption period. Amini's index estimates slightly

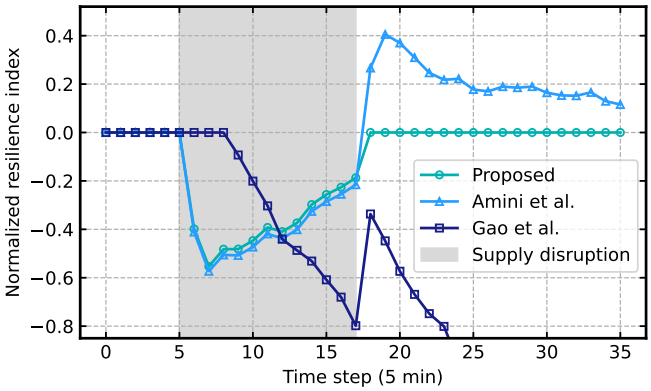


Figure 7.3: Comparison of MFD-based dynamic traffic resilience indicators (Munich LD-SSD scenario).

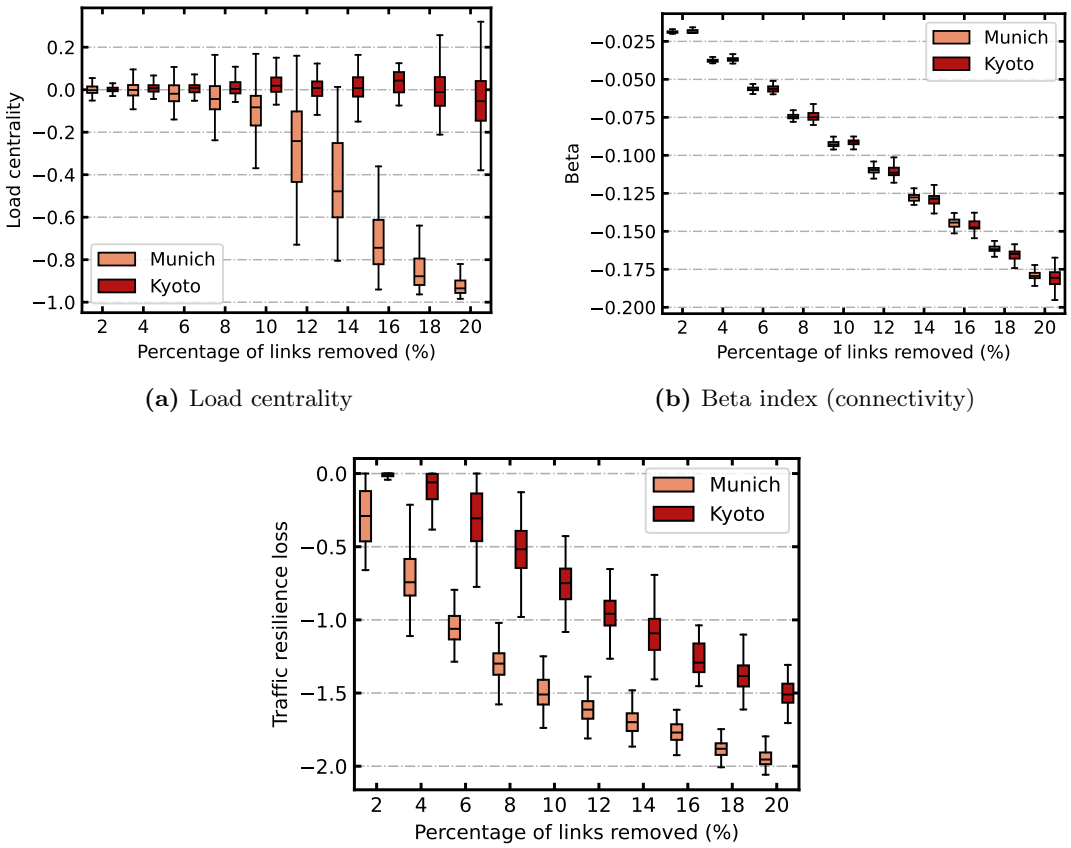
larger resilience loss during the disruption as its formulation neglects the reduction in network roadway length. Conversely, after the disruption, our index indicates the system recovers to the normal state, while Amini's index counts the process of serving the vehicles delayed by the disruption as an increase in resilience. In terms of Gao's index (Gao et al., 2022), no resilience loss is observed in the first 20 min of the disruption period. This can be attributed to the fact that, as highlighted in Section 2.3.2, this indicator was proposed for hyper-congestion, and cannot identify the resilience loss caused by supply-side disruptions. Consequently, their index, known as congestion deviation, fails to capture the degradation in traffic states accurately, leading to an incorrect estimation of traffic resilience loss during the disruption period. Afterward, Gao's index begins to decrease as a result of the rising number of vehicles in the network, where the accumulation surpasses its critical value. This index experiences a rebound of approximately 50% after the disruption as the system restores to its normal state and the critical accumulation also increases to its original value. It declines once more thereafter since the accumulation keeps increasing until the end of the investigation period. To provide a clearer presentation of the differences among these indices, Figure 7.3 does not display the entire curve of Gao's index. Its value steadily decreases in the invisible region.

7.3.3 Relationship between topological attributes and traffic resilience

This regression analysis uses topological variables as the explanatory dataset and resilience loss as the dependent variable. More specifically, the topological metrics listed in Table 7.2 of the network of each synthetic scenario are computed using the NetworkX (Hagberg et al., 2008) and OSMnx (Boeing, 2017) Python packages. The normalized resilience loss is estimated according to the approach introduced in Section 6.2.6 by using the traffic dynamics generated from SUMO. The lasso model described in Section 7.2.4 is adopted for the regression analysis, given its sparse output that can help identify the most significant topological attributes. A ρ value of 0.001 is used for the lasso model.

7.3.3.1 Explanatory analysis of selected variables

We found that variables describing the same topological attribute/property are most highly correlated. To address multicollinearity, we removed variables with a correlation coefficient exceeding 0.7 from the explanatory dataset, resulting in a retention of four variables: degree assortativity, load centrality, average edge length, and Beta index. Finally, the lasso model results in non-zero coefficients for the degree assortativity, load centrality, and Beta index for both cities, while the coefficient for average edge length is zero. Since the p-value for degree assortativity is very large (0.935 for Munich and 0.883 for Kyoto), indicating little significance, we also removed it from the final regression model.



(a) Load centrality

(b) Beta index (connectivity)

(c) Traffic resilience loss

Figure 7.4: Boxplots for the selected explanatory variables and the dependent variable.

Figure 7.4 shows the value of the remaining variables identified (i.e., load centrality and Beta index) and the dependent variable under different p values. Notably, each p value is associated with 100 scenarios, allowing us to construct boxplots that illustrate the distribution of variables. Different characteristics in load centrality can be observed

between the two cities. Specifically, load centrality in Kyoto hovers around 0 across different p values but decreases with the increase of p in Munich. Load centrality in Munich almost remains the same when $p \leq 0.08$, suggesting that local centrality might only be impacted significantly after a sufficient number of links are closed, a phenomenon that may also be present in Kyoto. On the contrary, the Beta index is distributed similarly in the two cities and has similar change patterns.

7.3.3.2 Regression model interpretation

Table 7.3 presents the model estimation results by solving the problem expressed by Equation (7.2). Only the variables with little correlations are inputted into the model, within which average edge length and degree assortativity are insignificant in both models, so it is not listed in the table. These two regression models reveal that significant topological attributes for traffic resilience are shared among different types of network topology, which is centrality and connectivity in this case. However, the significance levels and the extent to which they can explain traffic resilience are distinct. As suggested by the coefficients listed in Table 7.3, the Munich network is nine times and twice more sensitive than the Kyoto network in terms of centrality and connectivity, respectively.

Table 7.3: Regression model estimation.

Variable	Topology Attr.	Coef. [p-value] (Kyoto)	Coef. [p-value] (Munich)
Load centrality	Centrality	-0.1016 [0.25]	-0.9778 [<0.0001]
Beta index	Connectivity	8.1062 [<0.0001]	16.6719 [<0.0001]
Kyoto model		**Munich model**	
# of samples: 925		# of samples: 949	
R-squared: 0.8583		R-squared: 0.7894	

Recall that all explanatory variables have been transformed to be the percentage reduction compared to the complete network using Equation (7.1). No intercepts are estimated considering the sense of physical interpretation. Consequently, the coefficient of each specific variable reflects the sensitivity of traffic resilience to that variable. For instance, in the Kyoto model, a 10% reduction in the Beta index for three hours (the effective simulation duration) will lead to a loss of traffic resilience of about 0.81. According to the definition and interpretation of (normalized) traffic resilience provided in Section 6.2.5, it means that such a disruptive event can result in a cumulative service rate impairment of 81%. A larger absolute value of the coefficient for the Beta index in Munich suggests that Munich may experience more significant impairment than Kyoto when subjected to the same level of network connectivity reduction. This conclusion is also evident from Figure 7.4c, which demonstrates that with the same percentage of link removal, Munich exhibits a higher absolute traffic resilience loss compared to Kyoto. Furthermore, the relatively smaller coefficients for the other variables indicate that the Beta index is the most correlated and influential variable in our case stud-

ies. In other words, network connectivity plays a critical role in determining the traffic resilience of transportation networks. Conversely, the coefficient for load centrality in the regression models indicates that network centrality has a negative effect on traffic resilience. It means that urban planning should avoid increasing network centrality to enhance the network's traffic resilience to supply-side disruptions, which aligns with the findings from Wang et al. (2023) which found the negative relationship between node betweenness centrality and system resilience.

Further, referring back to Figure 7.4c, when the link removal probability p is small (≤ 0.04), the deviation of resilience loss in Munich is considerably larger than that in Kyoto. It implies that random link removal may cause more uncertainty in the traffic states in Munich under non-serious disruptions. This observation is further supported by the much greater R^2 obtained in the Kyoto model than in the Munich model. Therefore, it is fair to say that grid-like networks are more resilient and stable under supply-side disruptions in regard to traffic states. The good performance in R^2 validates that the proposed traffic resilience indicator based on MFD dynamics not only can represent traffic dynamics but also capture the network structure characteristics. This can also be seen from the high correlation between the Beta index and traffic resilience.

7.4 Conclusions and future work

In this chapter, we devote attention to the relationship between topological attributes and traffic resilience. To this end, we propose a synthetic supply-side disruption scenario generation procedure, which can be used to generate a diverse set of observations as input for a regularized linear regression model. This allows for the examination of various topological attributes and their impact on traffic resilience.

The experimental results provided evidence for the superiority of our proposed indicators over previous MFD-based indicators in terms of reliability and robustness. Our regression models for the two cities indicated that topological attributes significant to traffic resilience are generally consistent across different network topology. However, the extent to which these attributes explain traffic resilience differs. We found that network connectivity, measured by the Beta index in this study, emerged as the most correlated and significant attribute. A larger coefficient value for the Beta index in Munich suggests that Munich may experience a more pronounced service rate impairment than Kyoto when confronted with a similar reduction in network connectivity. This finding aligns with the conclusion that grid-like networks are more resilient than central ring networks offered in Chapter 6, which is also supported by previous literature (e.g., Zhang et al., 2015). The high correlation between traffic resilience and topological attributes proved that the proposed indicator based on MFD dynamics can not only represent traffic dynamics but also capture the network structure characteristics.

To further enhance our understanding of the relationship between traffic resilience and network topology, future research endeavors can explore additional types of network topology. Since this study primarily focused on two specific types of topology due to data constraints, it is important to acknowledge that the insights gained in this regard

are inherently limited. To overcome this limitation, one potential approach is to create synthetic networks representing various types of network topology using advanced traffic simulation models. A key challenge in synthetic network generation lies in the creation of reasonable OD demand matrices for comparative analysis. This study's findings suggest that a network's traffic resilience can be explained by its topological attributes to a certain extent. Therefore, integrating these attributes with other relevant information, such as network size and population data, can potentially serve as a surrogate for traffic resilience. The prospect of deriving a general relationship model from experiment results across numerous network scenarios is promising. Such a model can facilitate the evaluation of the transportation networks of real cities and assist in the design of resilient transportation networks in a simple way.

Chapter 8

Surrogate model for simulation-based traffic resilience optimization

The content of this chapter has been partially presented in the following work:

Lu, Q.-L., Sun, W., Lyu, C., Schmöcker, J.-D., and Antoniou, C. (2024d). Post-disruption lane reversal optimization with surrogate model to improve urban traffic resilience. *(Under Review)*.

8.1 Introduction and research contributions

The most common approach to transportation system resilience optimization problems has been the use of mathematical programming models. Stochastic programming (e.g., Chen and Miller-Hooks, 2012; Li et al., 2019; Zhang and Alipour, 2020; Alkhaleel et al., 2022) and bi-level optimization models (e.g., Faturechi and Miller-Hooks, 2014; Zhao and Zhang, 2020; Zou and Chen, 2021) have been extensively applied. These methods often incorporate a User Equilibrium (UE) component, such as the probabilistic UE (Lo and Tung, 2003), partial UE (Faturechi and Miller-Hooks, 2014), and elastic UE (Zhao and Zhang, 2020), as constraints for estimating travelers' response to optimization decisions. They can yield a macroscopic-level estimation of the influence of different decisions on system performance. Nonetheless, this estimation inevitably comes with a compromise in accuracy due to the strong assumptions (e.g., static travel demand) behind these simplified analytical UE components. This issue can be addressed by advanced traffic simulators that integrates disaggregate models of stochastic driver behaviors (e.g., car-following, lane-changing, route choice), which, once have been calibrated using real data, can provide accurate and detailed traffic state estimates for a given strategy (Osorio and Chong, 2015). However, traffic simulation models are generally computationally expensive due to their complex, dynamic, and stochastic nature. This violates another requirement of real-time optimization problems in the context of disruptions, i.e., being timely and efficient. Hence, it is natural to think of combining the advantages of traffic simulators in estimating time-dependent traffic dynamics and the efficiency of mathematical optimization. This inspiration has led to the development of surrogate-based simulation optimization algorithms.

This chapter proposes a surrogate model that couples a physical (problem-specific) component to provide structural information about the underlying problem and a func-

tional (general-purpose) component to improve the accuracy of the surrogate. In particular, we enhance the time-independent analytical network model proposed in Osorio (2019b) by adding a traffic state transition term in link-level density estimation, which results in a dynamic and temporally correlated model. Furthermore, a vectorized representation of the model is established to enable the application of an efficient Gradient Descent (GD) algorithm for variable calculation.

The contribution of this chapter is threefold as follows.

- We initialize a physics-informed surrogate model for simulation-based traffic resilience optimization problems by integrating a dynamic analytical network model and a regression model.
- We transform the analytical network model to a vectorized one to facilitate the application of efficient solution algorithms.
- A GD algorithm is derived to efficiently tackle the vectorized version of the proposed dynamic analytical network model.

The remainder of this chapter is organized as follows. Section 8.2 presents the proposed physics-informed surrogate model for traffic resilience optimization problems. Section 8.3 provides the solution algorithm for the surrogate model. Finally, Section 8.4 conclusions this section and points out future directions for research.

8.2 Surrogate model for simulation-based traffic resilience optimization

In this section, a physics-informed surrogate model is formulated for traffic simulators to enable the application of various optimization algorithms to determine well-performing decisions to improve traffic resilience with limited financial budgets and very tight computational budgets.

8.2.1 Physics-informed surrogate model and common constraints

For ease of description, we introduce the notation listed in Table 8.1. The variables with a subscript of τ indicate the interval-wise variant of the respective variables.

In fact, Equation 3.19 represents a physics-informed surrogate model, with the structural information on variables being captured by the physical component. In this section, we adapt the framework to the problems aiming at reducing traffic resilience loss. Considering traffic resilience optimization problems, performance indicators evaluating different decisions would be traffic resilience indicators. Here, we apply those proposed in Chapter 6, given their merits compared with others. The surrogate model for simulation-based traffic resilience optimization can then be expressed as

$$\text{(SOS-TR)} \min_{\mathbf{x}} \quad \mathscr{M} = \sum_{\tau=1}^{N_t} \eta_\tau R_\tau^A(\mathbf{x}_\tau, \mathbf{z}_\tau; \boldsymbol{\theta}) + \phi_\tau(\mathbf{x}_\tau; \boldsymbol{\beta}_\tau) \tag{8.1}$$

$$\text{s.t.} \quad R_\tau^A(\mathbf{x}_\tau, \mathbf{z}_\tau; \boldsymbol{\theta}) = \frac{1}{2}\left(\hat{\Delta}_\tau^s + \hat{\Delta}_{\tau-1}^s\right) T \qquad \tau = 1, \ldots, N_t \tag{8.2}$$

$$\hat{\Delta}_\tau^s = \max\left\{\hat{D}_\tau - \hat{D}_\tau^s, 0\right\} \qquad \tau = 1, \ldots, N_t \tag{8.3}$$

$$\hat{\Delta}_0^s = 0 \tag{8.4}$$

$$\hat{D}_\tau^s = \frac{1}{\gamma^s} \frac{\sum_{i \in \mathbb{L}} v_{i,\tau} k_{i,\tau} l_i}{\sum_{i \in \mathbb{L}} l_i} \qquad \tau = 1, \ldots, N_t \tag{8.5}$$

$$\mathbf{z}_\tau = h(\mathbf{x}_\tau, \mathbf{z}_{\tau-1}; \boldsymbol{\theta}) \qquad \tau = 1, \ldots, N_t \tag{8.6}$$

$$g(\mathbf{x}) \leq 0 \qquad \tau = 1, \ldots, N_t \tag{8.7}$$

Table 8.1: Notation in the surrogate model.

Notation	Description
τ	index of time intervals
i	index of links
r	index of routes
$\mathbf{x}$	decision variable of the surrogate optimization problem
$\mathbf{z}$	set of endogenous variables in the analytical network model
$\boldsymbol{\theta}$	set of exogenous variables in the analytical network model
$\boldsymbol{\beta}$	set of parameters of the functional component
$v_{i,\tau}$	space-mean speed of link i in time interval τ
$k_{i,\tau}$	density of link i in time interval τ
l_i	length of link i
$\lambda_{i,\tau}$	number of vehicles driving through link i in time interval τ
$\mathbf{v}_\tau$	vector of space-mean speed of all links in time interval τ
$\mathbf{k}_\tau$	vector of density of all links in time interval τ
$\boldsymbol{\lambda}_\tau$	vector of the number of vehicles driving through all links in τ
$h(\cdot)$	analytical network model

We replace $f^A(\mathbf{x}, \mathbf{z}; \boldsymbol{\theta})$ with $R^A(\mathbf{x}, \mathbf{z}; \boldsymbol{\theta})$ to demonstrate the focus on traffic resilience optimization. We consider a problem of N_t time intervals. The objective function is to minimize the estimated resilience loss over all intervals as expressed by Equation (8.1). Equations (8.2) and (8.3) calculate the resilience loss and the trip completion rate reduction in the case of discrete time intervals, respectively. Note, to ease modeling, we have replaced the original formulation of resilience loss with its absolute value here. Assuming that the period of interest starts at the beginning of the disruption, then we have $\hat{\Delta}_0^s = 0$ (Equation (8.4)). Equation (8.5) estimates the trip completion rate leveraging its linear relationship with the weighted space-mean flow, which was found in the empirical study conducted in Geroliminis and Daganzo (2008). Equation (8.6) represents the

analytical network model used to estimate traffic flow variables for deriving the physical component. In particular, $\mathbf{z}_\tau$ includes the estimation of space-mean speed, density, and arriving demand of all links, i.e., $\mathbf{z}_\tau = (\mathbf{v}_\tau, \mathbf{k}_\tau, \boldsymbol{\lambda}_\tau)$. The analytical network model in this study will be detailed later in Section 8.2.2. Finally, $g(\mathbf{x})$ is a vector function, representing a collection of practical constraints effected on the decision space. Examples of these constraints, to name just a few, include computational constraints (e.g., a limited number of simulation runs), financial constraints (e.g., limited repair budgets), personnel constraints (e.g., a limited number of repair crews), conservation constraints (e.g., the constant number of lanes of a two-directional road), positivity constraints, integer constraints.

To summarize, we propose a surrogate model (SOS-TR) combining a physical component and a functional component for approximating the unknown simulation-based objective function (Equation (3.18)) and modeling the difference between the approximation (after scaling) and simulation results, respectively. The constraints of the model include the analytical network model used to estimate the traffic flow situation under different decision strategies, the formulations to calculate the physical component using these analytical estimations, and other configuration constraints imposed on the decision space.

8.2.2 Dynamic and temporally correlated analytical network model

Table 8.2: Notation in the analytical network model.

Notation	Description
$P_{r,\tau}$	route choice probability of route r in time interval τ
$\mathbf{p}$	vector of route choice probabilities
O_r	the OD pair connected by route r
d_{O_r}	demand for OD pair O_r
$\mathbb{R}$	set of routes of the entire network
$\mathbb{R}_i^l$	set of routes containing link i
$\mathbb{R}_{O_r}^o$	set of routes connecting OD pair O_r
α_τ	value of time parameter in time interval τ
$t_{r,\tau}$	average travel time of route r in time interval τ
$t_{i,\tau}$	average travel time of link i in time interval τ
$n_{i,\tau}$	number of lanes of link i in time interval τ
$v_i^{\max}$	maximum speed (speed limit) of link i
k_i^{jam}	critical density (jam density) of link i
q_i^{cap}	capacity of link i
s_1, s_2	weighting factors of traffic states in two successive time intervals
N_{od}	number of OD pairs in the network
N_r	number of routes in the network
N_l	number of links in the network

The analytical network model developed in this study builds upon its time-independent predecessor proposed in Osorio (2019a). The latter model is described by a system of linear and nonlinear equations that simplifies the route choice model and the speed-density relationship model describing the traffic fundamental diagram. Although it uses time-dependent endogenous variables, the temporal correlation of traffic states is overlooked, i.e., the values of endogenous variables of different time intervals are independent of one another. To this end, we extend it to the temporally correlated case by adding a state transition term. We introduce the notation listed in Table 8.2.

The analytical network model requires as inputs OD demand matrices, from which the demand for available routes between OD pairs can be estimated leveraging a route choice model. Then, apply the flow propagation model that defines the relationship between route flows and link flows to transform route demand to link-level traffic flow states. The route choice model is defined on the travel costs of routes. This means link-level traffic flow states and route choice probabilities are interdependent. As a result, the analytical model adopted can be expressed as an equation system with intertwined variables. Specifically, the proposed analytical network model $\mathbf{z}_\tau = h(\mathbf{x}_\tau, \mathbf{z}_{\tau-1}; \boldsymbol{\theta})$ is defined as below.

$$\lambda_{i,\tau} = \sum_{r \in \mathbb{R}_i^l} P_{\tau,r} d_{O_r,\tau} \qquad \forall i \in \mathbb{L} \tag{8.8}$$

$$v_{i,\tau} = \frac{v_i^{\max}}{k_i^{\text{jam}}} \left(k_i^{\text{jam}} - k_{i,\tau} \right) \qquad \forall i \in \mathbb{L} \tag{8.9}$$

$$k_{i,\tau} = s_{1,\tau} \frac{k_i^{\text{jam}}}{q_i^{\text{cap}}} \frac{\lambda_{i,\tau}}{n_{i,\tau}} + s_{2,\tau} k_{i,\tau-1} \qquad \forall i \in \mathbb{L} \tag{8.10}$$

where the intermediate variables are given by

$$P_{r,\tau} = \frac{e^{\alpha_\tau t_{r,\tau}}}{\sum_{j \in \mathbb{R}_{O_r}^o} e^{\alpha_\tau t_{j,\tau}}} \qquad \forall r \in \mathbb{R} \tag{8.11}$$

$$t_{r,\tau} = \sum_{i \in \mathcal{L}(r)} t_{i,\tau} \qquad \forall r \in \mathbb{R} \tag{8.12}$$

$$t_{i,\tau} = \frac{l_i}{v_{i,\tau}} \qquad \forall i \in \mathbb{L} \tag{8.13}$$

Equation (8.8) estimates the expected arriving demand for link i by the sum of the products of the probability for routes containing link i and the expected OD demand of those routes. Equation (8.11) represents a route choice model by which vehicles opt for routes based on the expected travel time. α_τ indicates the value of time in τ. Equation (8.12) calculates the expected route travel time by simply aggregating the travel time of composing links, while Equation (8.13) estimates the expected link travel time. Equation (8.9) applies the Greenshield speed-density relationship model to specify link-level traffic flows. Equation (8.10) relates the link's expected density ($k_{i,\tau}$) to its expected demand ($\lambda_{i,\tau}$) as well as the density of the previous time interval ($k_{i,\tau-1}$). Considering

the interdependence of the endogenous variables in this model, Equation (8.10) can be recognized as the state transition model of traffic flows. $s_{1,\tau}$ and $s_{2,\tau}$ are scalars weighting the influence of the demand within the current time interval and the traffic state of the previous time interval. It is worth noting that in the system described above, endogenous variables include $\lambda_{i,\tau}$, $v_{i,\tau}$ and $k_{i,\tau}$, while exogenous variables include α_τ, $v_i^{\max}$, k_i^{jam}, q_i^{cap}, $s_{1,\tau}$ and $s_{1,\tau}$.

Compared to the original analytical model in Osorio (2019a), we have made modifications from three aspects: (i) A state transition term is added to the calculation of the expected density for capturing the temporal correlation of traffic states. (ii) The exogenous parameters irrelevant to links, i.e., α_τ, $s_{1,\tau}$ and $s_{1,\tau}$, are set to be time-dependent. We also develop an offline calibration approach to obtain plausible values for these parameters, which will be detailed in Section 8.3.3. (iii) We replace the Newell-type single-regime speed-density model therein with the linear Greenshield model to support the application of vectorization to further improve the computational efficiency of the network model. The model vectorization process will be detailed in Section 8.3.1. We would like to highlight that the proposed dynamic and temporally correlated analytical network model is expected to outperform its predecessor, particularly in scenarios involving varied network structures in different time intervals. This improvement in performance is attributed to the model's ability to capture the impact of network changes through the transition term in Equation (8.10). Hence, this model is very suitable for traffic resilience optimization problems, particularly in the evaluation of supply disruptions with (temporary) network damages or controls.

According to Equation (8.6), in addition to the traffic state from the previous time interval (here $\mathbf{k}_{\tau-1}$), outputs of the analytical network model are also dependent on the decision variable $\mathbf{x}$. However, $\mathbf{x}$ is not involved in the proposed network model. This is because $\mathbf{x}$ is usually defined to be related to the variables explaining the network structure or traffic flows involved in the network model. For instance, for a problem intending to optimize lane reversal control strategies, $n_{i,\tau}$ will be replaced with $x_{i,\tau}$. As a practical application, in Chapter 9, we will present a recovery measure optimization problem, in which the lane reversal technique is employed as the control decision.

8.3 Solution algorithms

In this section, we sequentially elucidate the vectorization of the analytical network model and its corresponding solution algorithm, and the calibration algorithm for the exogenous variables involved.

8.3.1 Vectorization of the analytical network Model

The dynamic analytical network model requires a scalable, efficient solution algorithm to enhance its practical application. Otherwise, opting for more computational resources to execute simulation evaluations may be more promising than solving the approximate subproblem inefficiently (Osorio, 2019a). Consequently, we want to vectorize the pro-

posed analytical network model to facilitate the development of a GD for solving required traffic flow variables.

Denote the number of OD pairs, the number of available routes, and the number of links in the network by N_{od}, N_r and N_l, respectively. The structure of a network can then be described by the following two matrices.

- Route-OD matrix: $\boldsymbol{B} = [b_{r,O}]_{N_r \times N_{od}}$, where $b_{r,O} = 1$ if route r connects OD pair O, otherwise 0.
- Route-link matrix: $\boldsymbol{C} = [c_{r,i}]_{N_r \times N_l}$, where $c_{r,i} = 1$ if link i is in route r, otherwise 0.

As an illustrative example, Figure 8.1 shows a simple directed network with 7 routes, 5 OD pairs (between nodes), and 5 links. The two network structure matrices for this network is given in Figure 8.2.

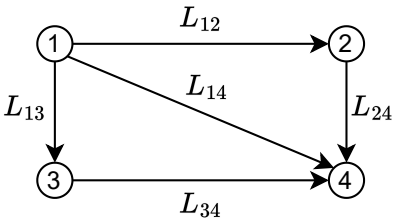


Figure 8.1: A simple example network.

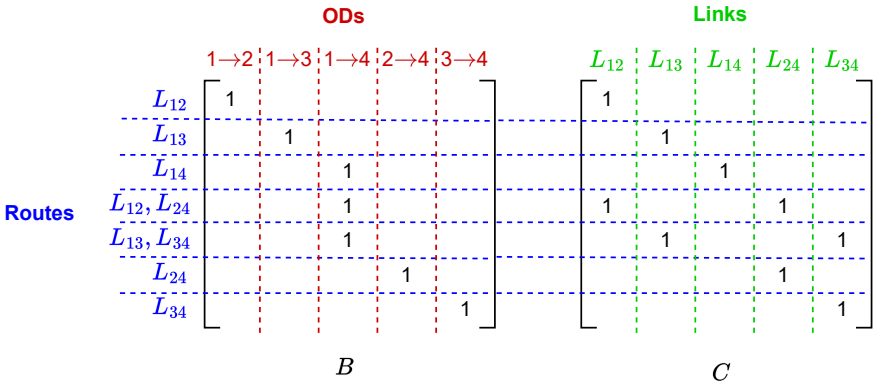


Figure 8.2: Examples of the route-OD matrix and the route-link matrix.

To facilitate the solution for the system of equations defined by Equation (8.8) - (8.10), we present the vectorized version of the analytical network model leveraging the above

network structure matrices as below.

$$\boldsymbol{\lambda} = \boldsymbol{C}^\top (\mathbf{p} \circ (\boldsymbol{B}\mathbf{d})) \tag{8.14}$$

$$\mathbf{p} = \exp[\alpha \boldsymbol{C}\mathbf{t}] \oslash \left(\boldsymbol{B}\boldsymbol{B}^\top \exp[\alpha \boldsymbol{C}\mathbf{t}]\right) \tag{8.15}$$

$$\mathbf{t} = \mathbf{l} \oslash \mathbf{v} \tag{8.16}$$

$$\mathbf{v} = \left(\mathbf{v}^{\max} \oslash \mathbf{k}^{\text{jam}}\right) \circ \left(\mathbf{k}^{\text{jam}} - \mathbf{k}\right) \tag{8.17}$$

$$\mathbf{k} = s_1 \left(\mathbf{k}^{\text{jam}} \oslash \mathbf{q}^{\text{cap}}\right) \circ (\boldsymbol{\lambda} \oslash \mathbf{n}) + s_2 \mathbf{k}^{(\text{p})} \tag{8.18}$$

where the lowercase letters in bold represent the vectors of the corresponding scalar variables listed in Table 8.1. $\mathbf{p}$ is an N_r-dimensional vector indicating the route choice probabilities. $\mathbf{d}$ is an N_{od}-dimensional vector, a flattened version of the OD matrix. The rest vectors are N_l-dimensional. Note, for simplicity of presentation, we have removed the index of time intervals τ from corresponding variables, and the expected density of the time interval $\tau - 1$ is indicated with a superscript (p), i.e., $\mathbf{k}^{(\text{p})}$. Originally, the system consists of six equations, but the vectorized version only contains five. This is because Equation (8.12) has been substituted into Equation (8.11) to derive Equation (8.15). Operations $\circ$ and $\oslash$ denote Hadamard (element-wise) product and division, respectively. $\exp[\cdot]$ denotes the element-wise exponential. $\boldsymbol{C}^\top$ indicates the transpose of matrix $\boldsymbol{C}$.

8.3.2 Solution algorithm for the analytical network Model

To calculate the endogenous variables, we construct the following problem

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & c \\ \text{subject to} \quad & \boldsymbol{\lambda} = \boldsymbol{C}^\top (\mathbf{p} \circ (\boldsymbol{B}\mathbf{d})) \end{aligned}$$

where c is an arbitrary constant such as 0. We can get the expected demand for all links $\boldsymbol{\lambda}^*$ by solving this problem. The expected density and speed can then be obtained by calculating Equation (8.18) and (8.17) sequentially. By applying the Lagrangian multiplier method, it can be transformed into the following problem

$$\text{(AN)} \quad \min_{\boldsymbol{\lambda}} \quad g = \frac{1}{2} \left(\boldsymbol{C}^\top (\mathbf{p} \circ (\boldsymbol{B}\mathbf{d})) - \boldsymbol{\lambda}\right)^\top \left(\boldsymbol{C}^\top (\mathbf{p} \circ (\boldsymbol{B}\mathbf{d})) - \boldsymbol{\lambda}\right) \tag{8.19}$$

Problem (AN) is an unconstrained problem and we can work out a GD algorithm to solve it. The gradient of g is given by

$$\nabla g = \left(\frac{\partial \boldsymbol{C}^\top (\mathbf{p} \circ (\boldsymbol{B}\mathbf{d}))}{\partial \boldsymbol{\lambda}} - \boldsymbol{I}_{N_l}\right) \left(\boldsymbol{C}^\top (\mathbf{p} \circ (\boldsymbol{B}\mathbf{d})) - \boldsymbol{\lambda}\right) \tag{8.20}$$

where $\boldsymbol{I}_{N_l}$ is an $N_l \times N_l$ identity matrix, and

$$\begin{aligned}\frac{\partial \boldsymbol{C}^\top(\mathbf{p}\circ(\boldsymbol{B}\mathbf{d}))}{\partial \boldsymbol{\lambda}} &= \boldsymbol{C}^\top \frac{\partial \mathbf{p}\circ\tilde{\mathbf{d}}}{\partial \boldsymbol{\lambda}} \\ &= \boldsymbol{C}^\top\left(\mathrm{diag}(\mathbf{p})\frac{\partial\tilde{\mathbf{d}}}{\partial\boldsymbol{\lambda}} + \mathrm{diag}(\tilde{\mathbf{d}})\frac{\partial\mathbf{p}}{\partial\boldsymbol{\lambda}}\right) \\ &= \boldsymbol{C}^\top\mathrm{diag}(\tilde{\mathbf{d}})\frac{\partial\mathbf{p}}{\partial\boldsymbol{\lambda}}\end{aligned}$$

where $\tilde{\mathbf{d}} = \boldsymbol{B}\mathbf{d}$ (independent of $\boldsymbol{\lambda}$) and the diag$(\cdot)$ operator represents the diagonal matrix constructed using the elements of the vector effected on. According to the chain rule, we can calculate

$$\frac{\partial\mathbf{p}}{\partial\boldsymbol{\lambda}} = \frac{\partial\exp[\alpha\boldsymbol{C}\mathbf{t}]\oslash\left(\boldsymbol{B}\boldsymbol{B}^\top\exp[\alpha\boldsymbol{C}\mathbf{t}]\right)}{\partial\boldsymbol{\lambda}} \tag{8.21}$$

$$= \left(\exp[\alpha\boldsymbol{C}\mathbf{t}][\circ](\alpha\boldsymbol{C}\frac{\partial\mathbf{t}}{\partial\boldsymbol{\lambda}})\right)[\oslash]\left(\boldsymbol{B}\boldsymbol{B}^\top\exp[\alpha\boldsymbol{C}\mathbf{t}]\right) - \tag{8.22}$$

$$\left(\exp[\alpha\boldsymbol{C}\mathbf{t}][\circ]\Big(\boldsymbol{B}\boldsymbol{B}^\top\big(\exp[\alpha\boldsymbol{C}\mathbf{t}][\circ](\alpha\boldsymbol{C}\frac{\partial\mathbf{t}}{\partial\boldsymbol{\lambda}})\big)\Big)\right)[\oslash]\left(\boldsymbol{B}\boldsymbol{B}^\top\exp[\alpha\boldsymbol{C}\mathbf{t}]\right)^{\circ 2} \tag{8.23}$$

where $(\cdot)^{\circ 2}$ denotes the Hadamard power, $[\circ]$ and $[\oslash]$ denote the penetrating face product and division, respectively. Considering a m-dimensional vector $\mathbf{y}$ and a m-by-n matrix $\boldsymbol{A}$, then $\mathbf{y}^{\circ 2} = [y_1^2, y_2^2, \ldots, y_m^2]^T$, and

$$\mathbf{y}[\circ]\boldsymbol{A} = \begin{bmatrix} y_1A_{11} & y_1A_{12} & \ldots & y_1A_{1n} \\ y_2A_{21} & y_2A_{22} & \ldots & y_2A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_mA_{m1} & y_mA_{m2} & \ldots & y_mA_{mn}\end{bmatrix}$$

$$\boldsymbol{A}[\oslash]\mathbf{y} = \begin{bmatrix} A_{11}/y_1 & A_{12}/y_1 & \ldots & A_{1n}/y_1 \\ A_{21}/y_2 & A_{22}/y_2 & \ldots & A_{2n}/y_2 \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}/y_m & A_{m2}/y_m & \ldots & A_{mn}/y_m\end{bmatrix}$$

Since both $\exp[\alpha\boldsymbol{C}\mathbf{t}]$ and $\left(\boldsymbol{B}\boldsymbol{B}^\top\exp[\alpha\boldsymbol{C}\mathbf{t}]\right)^{\circ 2}$ result in a N_r-dimensional vector, the calculation of $\frac{\partial\mathbf{p}}{\partial\boldsymbol{\lambda}}$ can be achieved using the penetrating face operations above.

Similarly, combining Equation (8.16), (8.17) and (8.18), we have

$$\frac{\partial\mathbf{t}}{\partial\boldsymbol{\lambda}} = \mathrm{diag}\left((s_1\mathbf{l}\circ\mathbf{v}^{\max})\oslash(\mathbf{v}^{\circ 2}\circ\mathbf{q}^{\mathrm{cap}}\circ\mathbf{n})\right) \tag{8.24}$$

which only contains exogenous variables. We can get $\frac{\partial\mathbf{p}}{\partial\boldsymbol{\lambda}}$ by substituting Equation (8.24) into Equation (8.23). Therefore, at each algorithm iteration, by conducting the calculations backward, we can obtain the gradient ∇g.

Finally, according to the GD algorithm, $\boldsymbol{\lambda}$ can be updated with the following rule:

$$\boldsymbol{\lambda} := \boldsymbol{\lambda} - \mu \nabla g \tag{8.25}$$

where μ is the step size. Due to the dependence on the previous interval, the vectorized system needs to be solved sequentially for each time interval. For convenience, further details on the procedure of the GD algorithm are presented in Algorithm 3.

Algorithm 3 Gradient Descent for solving the vectorized analytical network model.

For a given time interval τ and a given density vector for the previous interval $\mathbf{k}^{(p)}$.
Initialize exogenous variables α_τ, $s_{1,\tau}$ and $s_{2,\tau}$, $\forall \tau$.
Initialize the first guess $\boldsymbol{\lambda}^{(0)}$, algorithm step size μ, maximum iterations N_{max}.
Initialize the iteration index: $k = 0$.
for $k < N_{max}$ **do**
  Compute link-level traffic flows under $\boldsymbol{\lambda}^{(k)}$:
    – Compute the vector of link density $\mathbf{k}$ via Equation (8.18).
    – Compute the vector of speed $\mathbf{v}$ via Equation (8.17).
  Compute the partial derivative of the link travel time vector $\mathbf{t}$ with respect to $\boldsymbol{\lambda}$ via Equation (8.24).
  Compute the partial derivative of the route probability vector $\mathbf{p}$ with respect to $\boldsymbol{\lambda}$ via Equation (8.23).
  Compute the gradient direction $\nabla g^{(k)}$ via Equation (8.20).
  if $\nabla g^{(k)}$ does not change ($< \epsilon_{max}$) **then**
    Output the solution: $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^{(k)}$.
    break
  end if
  Improve the operation strategy: $\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} - \mu \nabla g^{(k)}$.
  Update the iteration index: $k = k + 1$.
end for

8.3.3 Surrogate model exogenous parameter calibration

The values of α_τ, $s_{1,\tau}$ and $s_{2,\tau}, \forall \tau$ directly influence the route choice behaviors and the relationship between demand and link performance. Applying appropriate values to these variables is important to the accuracy and reliability of the proposed analytical network model, i.e., the similarity of its results to simulation results. That is to say, these variables should be calibrated to make the analytical network model generates similar results to traffic simulations. Furthermore, considering that we are focusing on the MFD-based traffic resilience, the objective function of the calibration should be targeted on reducing the difference between the trip completion rates estimated from simulation results and analytical results. On the other hand, in practice, we are blind to the disruption data so the calibration can only be conducted offline by using the data under normal conditions. Even though the preference of drivers/vehicles might change after an occurrence of disruption, the variables calibrated via this method can

still provide some beneficial information to the approximation, which has also been proven by the experiment results.

Denote $\psi_\tau = \{\alpha_\tau, s_{1,\tau}, s_{2,\tau}\}$. Then, the offline calibration is to solve the following problem.

$$\text{(AN-C)} \quad \min_{\psi_\tau} \quad (\mathbf{v}_\tau \circ \mathbf{k}_\tau - \tilde{\mathbf{q}}_\tau)^\top (\mathbf{v}_\tau \circ \mathbf{k}_\tau - \tilde{\mathbf{q}}_\tau) \tag{8.26}$$

$$\text{subject to} \quad \text{Equation (8.14) - (8.18)} \tag{8.27}$$

where $\tilde{\mathbf{q}}_\tau$ indicates the vector of simulated traffic volumes in time interval τ. Clearly, this calibration model aims to reducing the difference between analytical and simulated traffic volumes. The problem can be solved via various derivative-free algorithms, such as the SPSA algorithm (Spall, 1998).

8.4 Conclusions and future work

Inspired by the strengths of traffic simulators in estimating time-dependent traffic dynamics and the efficiency of mathematical optimization, this chapter presents a physics-informed surrogate model for urban traffic resilience optimization problems. The surrogate model comprises a physical component derived from a deterministic dynamic analytical network model, coupled with a functional component for reducing the difference between simulated and analytical values. To further facilitate the application of the proposed dynamic and temporally correlated network model, we present its vectorized version and develop an efficient GD algorithm to solve the equation system. In order to improve the approximation performance of the analytical network model to traffic simulations, a calibration model is established to calibrate the exogenous variables in the network model offline using the simulation data under normal conditions.

We note that most of the computational cost of this surrogate model comes from the estimation of the route choice model defined as a Logit model here. Future research can be performed to simplify the route choice model or release the requirement for such a model to further improve the efficiency of the surrogate model. In addition, considering that potential network structure changes may lead to significant changes in the exogenous parameters involved in the analytical network model, the surrogate model can also benefit from an online calibration for these parameters.

Chapter 9

Simulation-based recovery measure optimization

The content of this chapter has been partially presented in the following work:

Lu, Q.-L., Sun, W., Lyu, C., Schmöcker, J.-D., and Antoniou, C. (2024d). Post-disruption lane reversal optimization with surrogate model to improve urban traffic resilience. *(Under Review)*.

9.1 Introduction and research contributions

This chapter presents an application of the surrogate modeling framework developed in Chapter 8. The application focuses on the optimization of recovery measures for strengthening urban traffic resilience in the immediate aftermath of supply disruptions.

Recall that system resilience is usually quantified by the integral of system functionality reductions over the entire disruption period. Implementing recovery measures can be useful for mitigating resilience loss through multiple ways such as increasing the recovery speed and/or improving the worst functionality state. In transportation systems, recovery measures are generally performed to reduce traffic congestion caused by disruptions. In terms of supply disruptions, several types of recovery measures have been investigated in the literature, including signal control, perimeter control, road expansion, and lane reversal control, etc.

The lane reversal technique (also referred to as contraflow) is adopted as the recovery measure. Lane reversal refers to the action of reversing the direction of lanes to reallocate the capacities of two-directional roads to accommodate the unbalanced traffic demand in the two directions (Zhang et al., 2019). Lane reversal has been implemented in some big cities to deal with traffic congestion issues during peak hours and large events (e.g., sport matches, concerts) (Zhao et al., 2014). It is a simple yet popular measure attributed to its flexibility without the need for advance planning, design, and construction of new lanes. To the best of our knowledge, this is the first attempt to realize this technique in traffic simulations for the sake of optimizing traffic resilience. According to the categorization method defined by Farahani et al. (2013), this problem can be classified as a urban road network design problem with operational decisions.

The contribution of this chapter is twofold as follows.

- We develop a physics-informed surrogate-based simulation optimization method to yield well-performing recovery measures, particularly lane reversal control strategies, with a very tight computational budget to improve traffic resilience for large-scale networks.
- We apply our approach to a real network in Kyoto, Japan, which has been calibrated using real observed data. The experiment results illustrate the practical value of lane reversal controls.

The remainder of this chapter is organized as follows. Section 9.2 presents the surrogate model for recovery measure optimization and develops solution algorithms for the problem. In Section 9.3, results from case studies are analyzed and discussed. Finally, Section 9.4 concludes this chapter.

9.2 Methodology

We introduce the notation listed in Table 9.1. The variables with a subscript of τ indicate the interval-wise version of the respective variables.

After the occurrence of a disruptive event, decision-makers have to respond quickly to propose an effective recovery strategy (a set of emergency recovery measures like traffic signals tuning, temporary lane reversals) to mitigate its negative impact and minimize social and economic losses. In this section, taking lane reversals as an example, we build a surrogate-based recovery measure optimization model to improve the transportation system traffic resilience to supply disruptions.

9.2.1 Definition of lane reversal control

By convention, for lane reversal control, decision variables are defined as binary variables indicating whether lane reversals are implemented on the corresponding lanes. For instance, for a two-directional road with two lanes each direction (shown in Figure 9.1a), decision variables can be expressed as

$$x_{ij} = \begin{cases} 1 & \text{if reverse the driving direction} \\ 0 & \text{otherwise} \end{cases} \tag{9.1}$$

where x_{ij} (i, $j \in [1,2]$) indicates the lane reversal decision on lane j of link i. In this case, we need to offer each lane a specific decision variable. The problem will be a 0-1 integer programming.

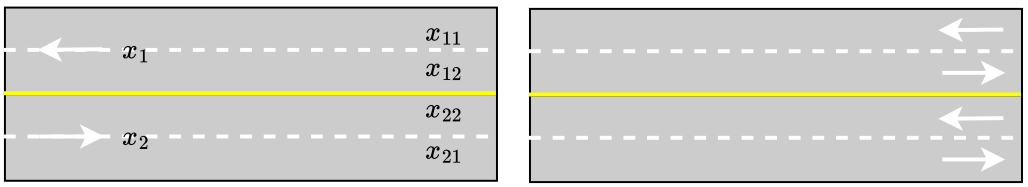
On the other hand, we can use the number of lanes in each direction as the decision variable, which can also reflect different lane reversal control situations. In this case, the decision will be:

$$x_i \in [1, 2, 3, 4] \tag{9.2}$$

where x_i ($i \in [1,2]$) represents the original number of lanes of link i. Clearly, such a definition can effectively reduce the dimension of decisions.

Table 9.1: Notation in the surrogate model for recovery measure optimization.

Notation	Description
$\mathbb{L}$	set of links in the network
$\mathbb{L}^c$	set of links allowed for implementing lane reversal
$\mathbb{L}^u$	set of links not allowed for implementing lane reversal
$\mathbf{x}$	decision variable of the surrogate optimization problem
$\mathbf{z}$	set of endogenous variables in the analytical network model
$\boldsymbol{\theta}$	set of exogenous variables in the analytical network model
$\boldsymbol{\beta}$	set of parameters of the functional component
$\boldsymbol{\delta}$	difference between the simulated and analytical resilience losses
$x_{i,\tau}$	number of lanes opened for link i in interval τ
$x_{i,\tau}^{(-)}$	number of lanes opened for the link opposite link i in τ
τ	index of time intervals
i	index of links
r	index of routes
n_i	original number of lanes of link i
$n_i^{(-)}$	original number of lanes of the link opposite link i
b^f	financial budget for implementing the recovery strategy
c_i^f	cost of implementing lane reversal on a lane of link i
$v_{i,\tau}$	space-mean speed of link i in time interval τ
$k_{i,\tau}$	density of link i in time interval τ
l_i	length of link i
$\lambda_{i,\tau}$	number of vehicles driving through link i in time interval τ
N_t	number of time intervals



(a) A two-directional road **(b)** An unacceptable decision

Figure 9.1: Lane reversal control (right-hand traffic).

It is noteworthy that in both definitions, combinations of lane reversal decisions resulting in alternating directions are typically not allowed to avoid confusion during implementation. For instance, the combination displayed in Figure 9.1b is generally prohibited. Therefore, appropriate constraints are needed for the first type of definition, while the second type of definition requires clarification that the lanes of each link are consecutive and adjacent to each other.

9.2.2 Simulation-based traffic resilience optimization

Here, we employ the second definition of lane reversal for the sake of reducing the dimension of decisions. Considering the cascading effect of traffic congestion, lane reversal decisions are distinguished for different time intervals. As a result, $\mathbf{x}$ has the same dimension as the number of links in the network times the time span (number of time intervals) of the control period. The surrogate model for recovery measure optimization can then be expressed as

$$\min_{\mathbf{x}} \quad \mathscr{M} = \sum_{\tau=1}^{N_t} \eta_\tau R_\tau^A(\mathbf{x}_\tau, \mathbf{z}_\tau; \boldsymbol{\theta}) + \phi_\tau(\mathbf{x}_\tau; \boldsymbol{\beta}_\tau) \tag{9.3}$$

$$\text{s.t.} \quad R_\tau^A(\mathbf{x}_\tau, \mathbf{z}_\tau; \boldsymbol{\theta}) = \frac{1}{2}\left(\hat{\Delta}_\tau^s + \hat{\Delta}_{\tau-1}^s\right) T \qquad \tau = 1, \dots, N_t \tag{9.4}$$

$$\hat{\Delta}_\tau^s = \max\left\{\hat{D}_\tau - \hat{D}_\tau^s, 0\right\} \qquad \tau = 1, \dots, N_t \tag{9.5}$$

$$\hat{\Delta}_0^s = 0 \tag{9.6}$$

$$\hat{D}_\tau^s = \frac{1}{\gamma^s} \frac{\sum_{i \in \mathbb{L}} v_{i,\tau} k_{i,\tau} l_i}{\sum_{i \in \mathbb{L}} l_i} \qquad \tau = 1, \dots, N_t \tag{9.7}$$

$$\mathbf{z}_\tau = h(\mathbf{x}_\tau, \mathbf{z}_{\tau-1}; \boldsymbol{\theta}) \qquad \tau = 1, \dots, N_t \tag{9.8}$$

$$\phi_\tau(\mathbf{x}_\tau; \boldsymbol{\beta}_\tau) \sim \mathcal{N}\left(\mu_\tau(\mathbf{x}_\tau), \mathbf{K}_\tau(\mathbf{x}_\tau)\right) \qquad \tau = 1, \dots, N_t \tag{9.9}$$

$$\frac{1}{2} \sum_{\tau=1}^{N_t} \sum_{i \in \mathbb{L}^c} c_i^f |n_i - x_{i,\tau}| \leq b^f \tag{9.10}$$

$$x_{i,\tau} + x_{i,\tau}^{(-)} = n_i + n_i^{(-)} \qquad \forall i \in \mathbb{L}, \tau = 1, \dots, N_t \tag{9.11}$$

$$x_{i,\tau} \geq 1 \qquad \forall i \in \mathbb{L}, \tau = 1, \dots, N_t \tag{9.12}$$

$$x_{i,\tau} = n_i \qquad \forall i \in \mathbb{L}^u, \tau = 1, \dots, N_t \tag{9.13}$$

We refer the reader to Section 8.2 for an explanation of Equations (9.3) - (9.8). Equation (9.9) represents the functional component, a GPR model. While many regression models (e.g., low-order polynomials and radial basis function) have been applied as a functional component in the literature, GPR is employed attributed to its proven ability to approximate the response surface for the transportation simulation input-output mapping (Chen et al., 2014; Beglerovic et al., 2017). More details about the GPR model development and fitting are given in Section 9.2.3.

Equation (9.10) represents the financial budget constraint. As the number of lanes of two opposite links are complementary, the total implementation cost is half of the sum of

the implementation costs of all links. Equation (9.11) expresses the conservation of the number of lanes of a two-directional road (i.e., a road with opposite links). Furthermore, here we assume lane reversal is not allowed to implement on all lanes of a link, implying that each link should have at least one lane after implementing the recovery strategy. This constraint is expressed as Equation (9.12). This maintains the integrity of routes, preventing route interruption and the resulting navigation failure that may cause confusion. Equation (9.13) says that the links that are not qualified for lane reversal should keep the number of opened lanes unchanged.

9.2.3 Functional component: Gaussian Process Regression

We define the resilience loss estimated based on simulation data as simulated resilience loss, and define the resilience loss derived from the analytical network model (after scaling) as analytical resilience loss. The functional component is a GPR model estimated using the simulated observations that relate the decision variables and the difference between simulated resilience loss and analytical resilience loss. To be specific, until iteration k, we obtain a dataset consisting of $N^{(k)}$ input vectors $\boldsymbol{X}^{(k)} = [\mathbf{x}_n]_{n=1}^{N^{(k)}}$ of dimension N_c (the number of links under control, i.e., $N_c = |\mathbb{L}^c|$), and the corresponding targets $\boldsymbol{\delta}^{(k)} = [\boldsymbol{\delta}_n]_{n=1}^{N^{(k)}}$. $\boldsymbol{X}^{(k)}$ is the matrix of recovery strategy samples that have been simulated until iteration k. $\boldsymbol{\delta}^{(k)}$ is the difference between the simulated and analytical resilience losses, i.e., $\boldsymbol{\delta}_n = \tilde{\boldsymbol{R}}^s(\mathbf{x}_n) - \boldsymbol{\eta}^{(k)} \circ \boldsymbol{R}^A(\mathbf{x}_n)$, where $\boldsymbol{\eta}^{(k)}$ indicates the vector of the scaling factors estimated at iteration k. $\mathbb{L}^c$ represents the set of controllable links.

The problem asks for a well-performing solution with a limited number of simulation observations. It is easy to cause the overfitting issue in high-dimensional GPR models, particularly in initial iterations, which may mislead the solution algorithm in an incorrect direction. To this end, although traffic states of a certain time interval are related to control decisions in previous intervals, we fit a GPR model for each time interval by using only the decisions of the corresponding interval (expressed by Equation (9.9)). In this way, the number of variables remains the same for all GPR models of different control intervals. However, this treatment may sacrifice some temporal correlation among control variables. Nevertheless, Osorio (2019a) demonstrated that an analytical network model, capable of capturing sufficient spatial correlation, can also incorporate temporal correlation in the surrogate model in a simple way. This is achieved by fitting the surrogate model with temporally correlated simulation observations. This approach aligns with our treatment described above. Moreover, in our model, the temporal correlation is further captured by the dynamic and temporally correlated analytical network model (Equation (8.6)) presented in Section 8.2.2.

For a certain time interval τ, we assume the true resilience loss difference $\mathbf{f}$ is a collection of random variables from a Gaussian process given by

$$\mathbf{f} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{K}) \tag{9.14}$$

where $\boldsymbol{\mu} = [m'(\mathbf{x}_1), m'(\mathbf{x}_2), \ldots, m'(\mathbf{x}_N)]$ and $\mathbf{K} = \mathbf{K}(\boldsymbol{X}, \boldsymbol{X})$ with $\mathbf{K}_{ij} = k'(\mathbf{x}_i, \mathbf{x}_j)$. m' represents the a priori mean function, and k' represents a positive definite kernel function. We set the a priori mean function $m'(\mathbf{x}_i) = \mathbf{0}, \forall i$. However, we have no access to true

function values ($\mathbf{f}$) in practical situations. In the k-th algorithm iteration, we denote the recovery measures that have been evaluated by $\boldsymbol{X}$ and denote the corresponding target by $\boldsymbol{\delta}$. Then, $\boldsymbol{\delta}$ is the observed value with noise introduced to true function values $\mathbf{f}$. We assume $\boldsymbol{\delta} = \mathbf{f} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$. As a result, $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathbf{K} + \sigma^2 \boldsymbol{I})$.

We can infer $\mathbf{f}_*$ given the strategies that have been evaluated ($\boldsymbol{X}$), corresponding target values ($\boldsymbol{\delta}$), and the strategies to be evaluated ($\boldsymbol{X}_*$). By definition, we have the joint distribution below

$$\begin{bmatrix} \boldsymbol{\delta} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{K} + \sigma^2 \boldsymbol{I} & \mathbf{K}_* \\ \mathbf{K}_*^\top & \mathbf{K}_{**} \end{bmatrix}\right) \tag{9.15}$$

where $\mathbf{K}_* = \mathbf{K}(\boldsymbol{X}, \boldsymbol{X}_*)$ and $\mathbf{K}_{**} = \mathbf{K}(\boldsymbol{X}_*, \boldsymbol{X}_*)$. As per Rasmussen and Williams (2008), by deriving the conditional distribution, the predictive equation for GPR can be expressed as

$$\bar{\mathbf{f}}_* = \mathbf{K}_*^\top [\mathbf{K} + \sigma^2 \boldsymbol{I}]^{-1} \boldsymbol{\delta} \tag{9.16}$$

Since we only utilize the mean function calculated by the posterior distribution of possible functions for regression predictions, the equation for calculating the covariance is omitted here.

9.2.4 Solution algorithm for the surrogate-based recovery measure optimization problem

The traffic resilience optimization problem in this study is a discrete simulation-based optimization problem. Considering the inclusion of the analytical network model consisting of implicit nonlinear equations in the problem, we integrate a Genetic Algorithm (GA) into the metaAHA framework present in Zhou et al. (2023), which was focused on car-sharing service network design problems, to solve this problem. AHA is for the Adaptive Hyperbox Algorithm proposed in Xu et al. (2013), which was developed for solving high-dimensional discrete optimization via simulation problems. However, due to the limited search space of each decision variable in our case (for a two-directional road with two lanes each direction, the decision space is $\{1, 2, 3\}$), defining and updating the hyperbox at each approximating state becomes unnecessary since all decision combinations can be regarded as within the neighbourhood of each other. Therefore, rather than relying on the hyperbox to perform exploitation, we make use of the best solutions of the last N_{ga} epochs in GA. Furthermore, we also employ random sampling to undertake the exploration task for this algorithm, specifically generating N_{rs} random samples in each simulation iteration. Regarding the computational budget, we follow a similar setup to that in Osorio and Chong (2015), i.e., the tight computational budget is defined as a maximum number of simulation runs that can be carried out. We denote it by N. Algorithm 4 presents the algorithm adopted to solve the proposed simulation-based traffic resilience optimization problem. A flowchart of the algorithm is provided in Figure 9.2 for clarity and ease of understanding.

Algorithm 4 Simulation-based traffic resilience optimization.

1: **Step 0: Initialization**

- Initialize analytical network model
  - Set α_τ, $s_{1,\tau}$, $s_{2,\tau}, \forall\tau$ according to the off-line calibration results.
  - Initialize other exogenous variables $\mathbf{d}_\tau, \forall\tau, \mathbf{v}^{\max}, \mathbf{k}^{\text{jam}}, \mathbf{q}^{\text{cap}}$.
  - Define the network structure matrices $\boldsymbol{B}$ and $\boldsymbol{C}$.
  - Initialize the density of the interval previous to the period of interest: $\mathbf{k}_1^{(\text{p})} = \mathbf{k}_0$.
- Set the maximum number of iteration runs N. Set the number of strategies resulting from GA to preserve N_{ga}, the number of strategies generated by random sampling N_{rs} at each iteration.
- Initialize the iteration index $k = 1$. Current number of iteration runs $N^{(k)} = 0$.

2: **Step 1: Strategy optimization and sampling**

- Apply GA to solve SOS-TR. Keep the best strategies of the last N_{ga} epochs: $\mathbb{X}_{ga}^{(k)}$.
- Randomly sample N_{rs} strategies in the decision space: $\mathbb{X}_{rs}^{(k)}$.
- Strategies to examine: $\mathbb{X}^{(k)} = \mathbb{X}_{ga}^{(k)} \cup \mathbb{X}_{rs}^{(k)}$.

3: **Step 2: Strategy evaluation**

- Run simulation with the strategies in $\mathbb{X}^{(k)}$.
- Evaluate the traffic resilience losses for all strategies: $\{|\tilde{R}^s(\mathbf{x})||\mathbf{x} \in \mathbb{X}^{(k)}\}$.
- Update the current best strategy: $\mathbf{x}^{*(k)} = \text{argmin}_{\mathbf{x}} |\tilde{R}^s(\mathbf{x})|$.
- Update iteration runs: $N^{(k)} = N^{(k)} + N_{ga} + N_{rs}$.

4: **Step 3: Termination check**

- If $N^{(k)} \geq N$, stop, otherwise, proceed to Step 4.

5: **Step 4: Surrogate model update**

- Evaluate the strategies in $\mathbb{X}^{(k)}$ using the surrogate model.
- Use all strategies evaluated so far to fit the scaling factor $\eta_\tau^{(k)}$ and the parameters in the functional component $\boldsymbol{\beta}_\tau^{(k)}$ for each time interval separately.
- Update the surrogate model: $\mathscr{M}^{(k)} = \sum_{\tau=1}^{N_t} \eta_\tau^{(k)} R_\tau^A(\mathbf{x}_\tau, \mathbf{z}_\tau; \boldsymbol{\theta}) + \phi_\tau^{(k)}(\mathbf{x}_\tau; \boldsymbol{\beta}_\tau^{(k)})$
- Set $k = k + 1$, proceed to Step 1.

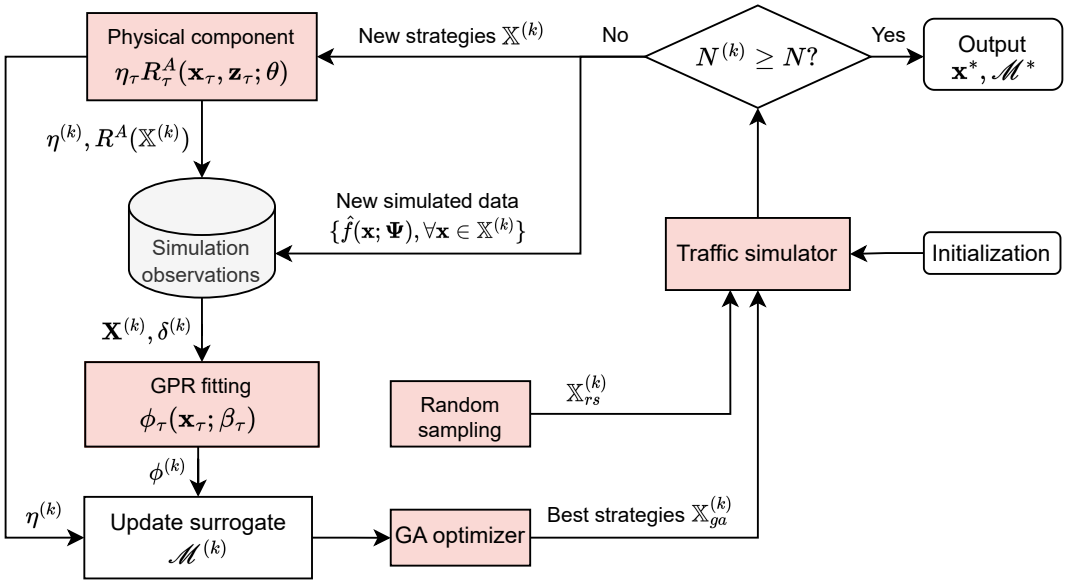


Figure 9.2: Flowchart for the solution algorithm.

9.3 Case study and results

We conduct case studies in the city center of Kyoto, Japan, which has been used in previous chapters, to examine the performance of the proposed surrogate model. We refer the reader to Section 3.4.2.2 for more details about the network structure and information on this study area. This network is a large-scale real network compared to those used in related literature (e.g., the 122-link freight transportation network in Li et al., 2019). Our analysis focuses on the morning period from 5 am to 12 pm, with the initial and final hours as warm-up and dissipation periods, respectively.

The supply disruption represented by temporary closing the links within the read area highlighted in Figure 3.8c is designed for experiments. The disruption period lasts from 6 am to 11 am. This scenario design follows similar styles as those presented in relevant studies (e.g., Li et al., 2019; Liao et al., 2018).

9.3.1 Simulation setup

SUMO is employed for conducting simulation evaluations under various recovery strategies. All simulations are conducted at the mesoscopic resolution with a non-iterative dynamic stochastic user assignment method to approximate the DUE. The mesoscopic model has been calibrated for morning-period demand using real observed data.

9.3.2 Analytical network model setup

This large-scale network provides tens of thousands of routes for vehicles to finish their trips. To balance algorithm efficiency without compromising the accuracy of the ana-

lytical network model, we filter out routes shorter than five minutes and with a demand fewer than five trips. This step leads to 188 effective zonal OD pairs, 3,836 effective routes and 1,023 effective links. Rather than a compromise, we treat it as a necessary step to reduce the influence of negligible OD pairs and uncertainty in route choice behaviors inherent in simulations. Therefore, the dimensions of the network structure matrices introduced in Section 8.3.2 are $N_{od} = 188$, $N_r = 3836$ and $N_l = 1023$.

Note, some links becomes unavailable in the link closure scenario described above. With the same filter, we obtain 110 effective OD pairs, 2,802 effective routes, and 890 effective links, denoted as $N_{od}^{\text{lc}} = 110$, $N_r^{\text{lc}} = 2802$ and $N_l^{\text{lc}} = 890$ with "lc" indicating link closure. Route-OD matrices ($\boldsymbol{B}$ and $\boldsymbol{B}^{\text{lc}}$) and route-link matrices ($\boldsymbol{C}$ and $\boldsymbol{C}^{\text{lc}}$) can then be constructed correspondingly. It is worth mentioning that these matrices are very sparse, and thus can be handled very efficiently with specific functions available in most programming languages, such as the *sparse* function in the *scipy* package for Python.

9.3.3 Link reversal control setup

In terms of the control period, we assume that the traffic control center needs one hour to make decisions and allocate necessary resources (e.g., personnel, control equipment, and materials) after the occurrence of the disruption. Therefore, the control period for all scenarios is from 7 am to 11 am, with each controlled time interval lasting one hour.

We employ $N = 150$, same as Osorio and Chong (2015), as the limit of simulation runs in Algorithm 4 to devise a contraflow-based recovery strategy. Set $N_{ga} = 10$ and $N_{rs} = 20$ for algorithm exploitation and exploration respectively. On the other hand, instead of constraining the problem with a financial budget as expressed in Equation (9.10), we specify a fix number of controlled links (N_c) for each time interval, representing the situation of limited personnel resources. Here, we consider three configurations for N_c, specifically $N_c \in \{10, 15, 20\}$. However, it is flexible to change without much efforts. Likewise, this should not influence the conclusions drawn from the subsequent evaluation and analysis of experimental results.

Regarding the selection of controlled links, while decision-makers can base their choices on local conditions, it is natural to have the following principles: (i) The road condition allows for implementing lane reversal; (ii) Selected roads exhibit imbalanced saturation rates in the two directions; (iii) One direction has a high saturation rate (i.e., the link is congested), as lane reversals will not significantly impact traffic states otherwise. More specifically, we consider links with a saturation rate ratio greater than 1.5 (principle ii) and a saturation rate greater than 0.75 (principle iii). We choose the N_c links with the most imbalanced saturation rates (largest ratios of saturation rates) for control.

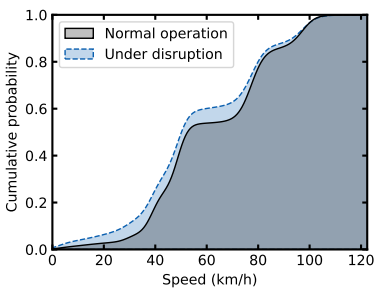
9.3.4 Traffic resilience evaluation

It is valuable to comprehend the resilience pattern of the Kyoto system under the disruption scenario presented in Section 3.4.2.2. This scenario will also be used to assess the proposed surrogate modeling approach in the following experiments and evaluations.

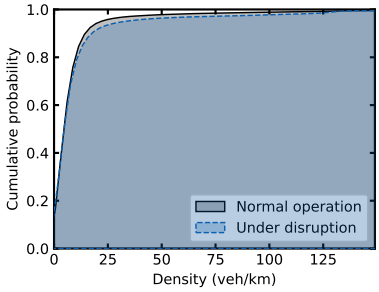
Figure 9.3a shows the cumulative distribution of link speeds during the disruption period (i.e., 6 am - 10 am). The difference between the normal and disruption operation conditions mainly happens within the speed ranges of [0, 40] and [50, 70]. About 5% of additional links observe a speed smaller than 40 km/h under the disruption condition compared to the regular operation. This also applies to the speed range of [50, 70]. It implies that link speeds are reduced due to the additional demand transferred from the routes blocked due to disruption. Correspondingly, the link densities observe an increase as presented in Figure 9.3b.

Figure 9.3c illustrates the change in trip completion rate, the system functionality measurement utilized in this study. Each point on the line represents the traffic state of a five-minute interval. The disruption period is indicated in grey. The total resilience loss is obtained by integrating trip completion rate reduction across the investigation period, represented as the shaded area in the figure.

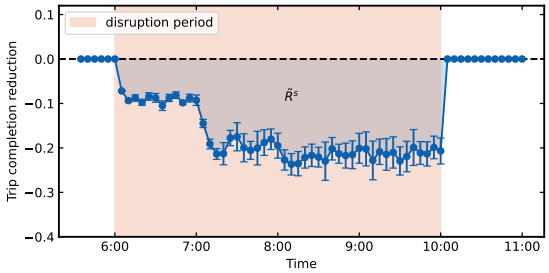
It is evident that the extent of completion rate reduction changes with the demand level. Compared to the normal situation, the reduction is around 10% from 6 am to 7 am, during which travel demand is relatively small. This value increases to about 20% later due to the rising demand and reaches a new degraded equilibrium. Attributed to its grid layout, the system offers relatively more route alternatives for vehicles to go to their destinations compared to networks with other types of layouts, such as ring networks. This characteristic, validated in Lu et al. (2024c), makes systems with grid networks more stable against disturbances, facilitating the formation of a new equilibrium.



(a) Speed cumulative distribution



(b) Density cumulative distribution



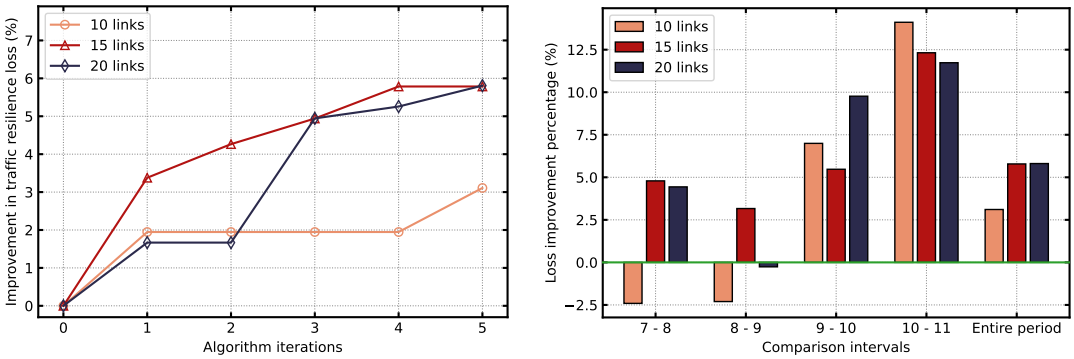
(c) Traffic resilience loss

Figure 9.3: Traffic resilience loss evaluation and comparison of traffic variables.

9.3.5 Algorithm performance in different control scenarios

Figure 9.4a shows the overall performance of the proposed algorithm under different control scenarios outlined in Section 9.3.3. We set the maximum simulation runs $N = 150$, and strategy evaluations for each iteration for exploitation $N_{ga} = 10$ (strategies recommended by GA) and exploration $N_{rs} = 20$ (strategies generated by random sampling). As a result, the algorithm terminates after five iterations. The x-axis in Figure 9.4a represents the iteration index, and the y-axis represents the percentage improvement in traffic resilience loss, which is given by $100(\tilde{R}^{s,\prime} - \tilde{R}^{s})/|\tilde{R}^{s}|$, where $\tilde{R}^{s,\prime}$ denotes the normalized resilience loss after implementing the recovery strategy. Clearly, the 15-link and 20-link control scenarios achieve similar resilience improvements after the algorithm terminates, which is about double of the one seen with the 10-link control scenario. Furthermore, it is fair to say that the 15-link scenario is preferable in this case study compared to the 20-link scenario for two main reasons: (i) The 15-link scenario achieves comparable efficacy to the 20-link scenario with fewer controlled links, making it more economical; (ii) The algorithm is more robust in the 15-link scenario, represented by its superior performance over the 20-link scenario in the first two iterations. Surprisingly, the algorithm in the 20-link scenario performs even worse than the 10-link scenario in the first two iterations. It is worth noting that these findings are specific to this case study. However, they highlight the importance of determining the number of controllable links in lane reversal control applications. Insufficient controlled links may lead to limited improvement in resilience loss, while excessive controlled links may increase implementation costs and reduce algorithm robustness. Consequently, selecting an appropriate number of controlled links should consider resilience improvement, implementation cost, and algorithm reliability. On the other hand, converting the problem with a maximum personnel constraint to a problem with a maximum budget constraint could help the algorithm automatically determine the optimal number of links given a predefined budget.

Figure 9.4b provides a comparison of the performance of the best solutions for the three control scenarios in different time intervals. It is interesting to see that in the 10-link scenario, the best solution leads to a reduction in trip completion rates during the 7 am - 8 am and 8 am - 9 am intervals. However, it results in the greatest percentage improvement in the 10 am - 11 am interval. The observation suggests that the best control strategy might not uniformly enhance traffic across the entire control period. Instead, it may sacrifice performance in certain periods to achieve better overall results. This phenomenon, on the other hand, reflects the temporal correlation of traffic states and the cascading effect of traffic congestion. Yet, the underlying reasons for these variations (e.g., the transfer of traffic congestion) would require further investigation beyond the scope of this study. Similar phenomenon is also observed in the 20-link scenario where the trip completion rate in interval 8 am - 9 am also sees a decline. In contrast, the best solution for the 15-link scenario demonstrates an ability to improve traffic resilience across all controlled time intervals, which provides additional facet representing the superiority of the 15-link scenario.



(a) Overall comparison (5 iterations) **(b)** Hourly comparison (5 iterations)

Figure 9.4: Algorithm performance evaluation and comparison (5 iterations).

9.3.6 Accuracy analysis of surrogate model

In addition to evaluating the overall effectiveness of the proposed approach in enhancing traffic resilience, it is crucial to analyze how the surrogate model evolves with algorithm iterations to fully uncover the factors influencing the performance of the approach. Figure 9.5 depicts the distributions for the difference between simulated resilience loss and approximated resilience loss of five algorithm iterations for three control scenarios and four time intervals. Violin plots are employed to visualize the distributions. Figure 9.5 offers three important points regarding the characteristics of the proposed surrogate model.

First, focusing on a single plot in the figure, the loss difference is distributed around zero, implying that the surrogate model provides an unbiased approximation to traffic simulation. Moreover, the reduction in loss difference with the progression of algorithm iterations suggests an increasingly accurate approximation. This pattern holds across all control scenarios and time intervals, except for the "15 links, 9 am - 10 am" scenario, where the distribution variance increases slightly in the 4th and 5th iterations compared to the previous two iterations.

Second, considering each row of plots in Figure 9.5, the variance of the loss difference increases with time, reflecting the enduring effect of previous controls on subsequent time intervals. Noteworthy, we only use the control variables of the corresponding time interval to fit the functional component for constructing the surrogate model. This means the influence of the controls in previous time intervals is not counted. While this serves as a compromise to mitigate the dimensionality of GP models and prevent overfitting, it also underscores the opportunity for enhancing the surrogate model through integrating effective dimension reduction techniques.

Third, for each column of plots in Figure 9.5, it can be observed that the variance of the loss difference increases with the number of controlled links. Similarly, an increased number of controlled links introduces a higher dimensionality to the GP models, poten-

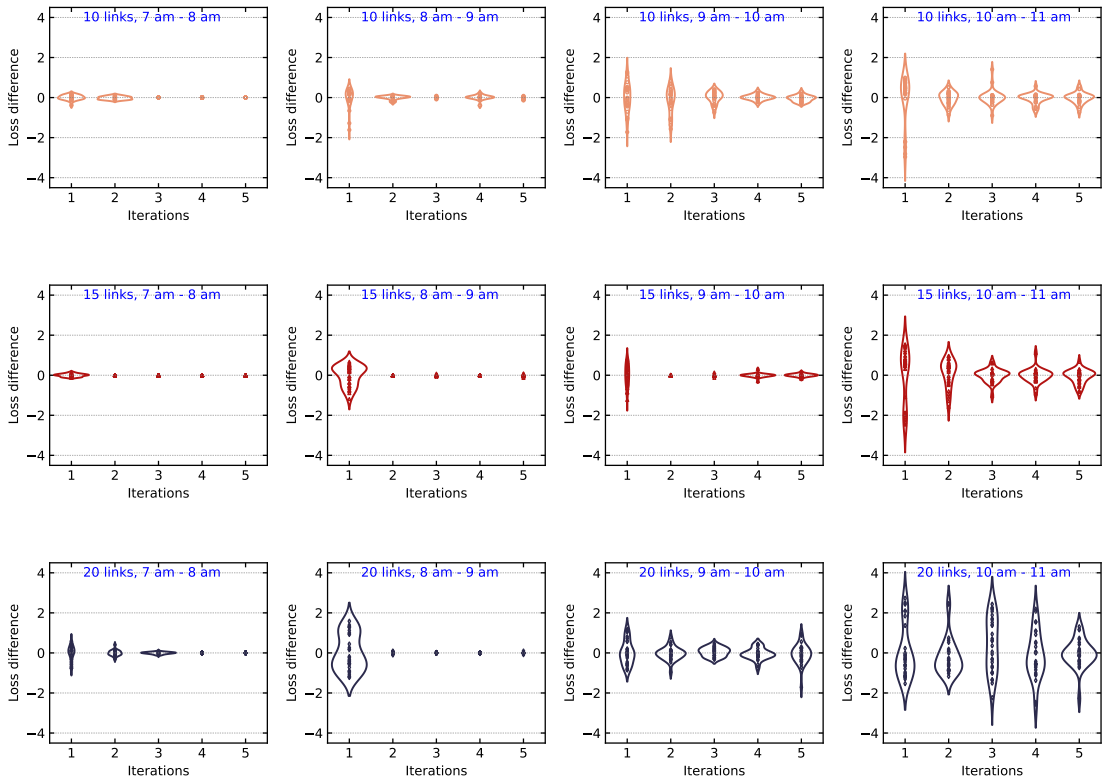


Figure 9.5: Difference between simulated loss estimate and analytical loss estimate.

tially leading to overfitting. This suggests that selecting representative variables to fit the GP could improve the approximation accuracy of the surrogate model.

9.3.7 Surrogate model structure validation

The proposed surrogate model consists of two main components: a physical component that approximates simulation results using an analytical network model, and a functional component that estimates the difference between simulated and analytical results (after appropriate scaling). In particular, GPR is applied as the functional component. Essentially, this model structure can be viewed as a physics-informed surrogate model. To demonstrate the merits of this model structure, it is crucial to examine the necessity of the physical component and explore the form of the function component.

Table 9.2: Role of different components in the surrogate.

Func. component	$\boldsymbol{\eta}$	s_2	# links	Optimal loss	% improvement
GPR ($\mathbf{x}$)	Fitted	Calibrated	10	-0.7205	3.11%
GPR ($\mathbf{x}$)	Fitted	Calibrated	15	-0.7006	5.79%
GPR ($\mathbf{x}$)	Fitted	Calibrated	20	-0.7004	5.81%
GPR ($\mathbf{x}$)	$[1]_{N_t}$	Calibrated	10	-0.7205	3.11%
GPR ($\mathbf{x}$)	$[1]_{N_t}$	Calibrated	15	-0.7069	4.94%
GPR ($\mathbf{x}$)	$[1]_{N_t}$	Calibrated	20	-0.7130	4.12%
GPR ($\mathbf{x}$)	$[0]_{N_t}$	Calibrated	10	-0.7291	1.94%
GPR ($\mathbf{x}$)	$[0]_{N_t}$	Calibrated	15	-0.7069	4.94%
GPR ($\mathbf{x}$)	$[0]_{N_t}$	Calibrated	20	-0.7205	3.11%
EN ($\mathbf{x}$)	Fitted	Calibrated	10	-0.7205	3.11%
EN ($\mathbf{x}$)	Fitted	Calibrated	15	-0.7069	4.94%
EN ($\mathbf{x}$)	Fitted	Calibrated	20	-0.7205	3.11%
EN ($\mathbf{x}, \mathbf{x}^{\circ 2}$)	Fitted	Calibrated	10	-0.7205	3.11%
EN ($\mathbf{x}, \mathbf{x}^{\circ 2}$)	Fitted	Calibrated	15	-0.7035	5.4%
EN ($\mathbf{x}, \mathbf{x}^{\circ 2}$)	Fitted	Calibrated	20	-0.7312	1.67%
GPR ($\mathbf{x}$)	Fitted	$[0]_{N_t}$	10	-0.7205	3.11%
GPR ($\mathbf{x}$)	Fitted	$[0]_{N_t}$	15	-0.7185	3.38%
GPR ($\mathbf{x}$)	Fitted	$[0]_{N_t}$	20	-0.7312	1.67%
Baseline (no control):				-0.7436	-

The necessity of the physical component can be evaluated from two perspectives, including the effect of the scaling factor $\boldsymbol{\eta}$ and the effect of the physical information derived from the analytical network model. Hence, two additional experiments associated with $\boldsymbol{\eta} = 1$ and $\boldsymbol{\eta} = 0$ (see Equation (9.3)) are designed and executed for each control scenario. Table 9.2 compares the percentage improvement in traffic resilience resulting from the best solutions for these experiments. We can see that setting $\boldsymbol{\eta} = 1$ leads to either the same or degrading performance in all control scenarios. Notably,

setting $\boldsymbol{\eta} = 1$ is equivalent to ignoring the systematic approximation bias of the analytical network model. The degrading performance indicates that the functional component cannot effectively capture such systematic bias with limited observations. Furthermore, the performance continues to decline when setting $\boldsymbol{\eta} = 0$, i.e., discarding the physical component completely from the surrogate model. In this case, the functional component fails to reproduce similarly accurate estimations as the complete version. In other words, the surrogate model informed by physical laws among traffic variables can guarantee better recovery strategies with the same simulation observations than the one without. To summarize, both the physical relationship capturing the correlation among traffic variables and the scaling factor accounting for the approximation bias of the analytical model play a vital role in the performance of the proposed surrogate model for recovery measure optimization.

Further, we validate the choice of GPR model as the functional component by comparing its performance with the simple Elastic Net (EN) model adopted in Osorio and Chong (2015) for traffic signal control problems. The EN model is a regularized linear regression model, which is defined as $\phi^{\mathrm{EN}}(\mathbf{x}) = \boldsymbol{\beta}^{\mathrm{EN}}\mathbf{x}$ where

$$\boldsymbol{\beta}^{\mathrm{EN}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \quad \|\boldsymbol{\delta} - \boldsymbol{X}\boldsymbol{\beta}\|^2 + \alpha_1\|\boldsymbol{\beta}\|_1 + \alpha_2\|\boldsymbol{\beta}\|^2.$$

α_1 and α_2 are hyperparameters. For comparative analysis, we use $\mathbf{x}_\tau$ to fit the EN model for the same interval. Moreover, to incorporate nonlinear features into the model, we also evaluate the model with $\mathbf{x}_\tau$ and $\mathbf{x}_\tau^{\circ 2}$ as inputs. The experiment results are also summarized in Table 9.2 (with subscript τ omitted). Clearly, the surrogate model employing GPR ($\mathbf{x}$) outperforms both the one with EN ($\mathbf{x}$) and the one with EN ($\mathbf{x}, \mathbf{x}^{\circ 2}$) across all control scenarios regarding the improvement in traffic resilience. This suggests that, at the approximating state of each algorithm iteration, the GPR model can more precisely describe the shape of the real objective function (i.e., traffic simulations) surface in the neighborhood of the current solution. As a result, the recovery strategies suggested by GPR are closer to the optimum. This aligns with the findings from Beglerovic et al. (2017), where GPR (Kriging) requires fewer system evaluations to yield the best results. Chen et al. (2014) also indicated that the surrogates with low-order polynomials may lead to biased estimations, particularly in highly nonlinear functions. On the other hand, by comparing the EN models, we note that EN ($\mathbf{x}, \mathbf{x}^{\circ 2}$) performs better than EN ($\mathbf{x}$) in the 15-link control scenario but exhibits worse performance in the 20-link control scenario. The potential explanation for the former situation is that incorporating the nonlinear features introduced by $\mathbf{x}^{\circ 2}$ enhances the approximation performance of the surrogate model. On the contrary, the latter situation might arise from incorporating an excessive number of variables in the functional component, leading to overfitting issues. This would mislead the search direction of the algorithm, consistent with the findings presented in Section 9.3.6.

Additionally, the benefits of the proposed dynamic and temporally correlated analytical network model over its predecessor are assessed by setting $s_2 = 0$. This adjustment renders the traffic transition term, incorporated into the density estimation function as expressed in Equations (8.10) and (8.18), inactive. Consequently, the proposed model re-

verts to the original static version. The comparison presented in Table 9.2 demonstrates that the surrogate with the static analytical network model exhibits inferior performance compared to those employing the dynamic model across all control scenarios.

9.4 Conclusions and future work

This chapter investigated the optimization problem of urban traffic resilience, focusing on emergency recovery measures, particularly lane reversal control in the aftermath of supply disruptions. A surrogate-based simulation optimization model was applied to address the problem. The surrogate model comprises a physical component derived from a deterministic dynamic analytical network model, coupled with a functional component driven by a GPR model. An algorithm integrating Gradient Descent and genetic search within the Adaptive Hyperbox Algorithm (AHA) framework was developed to solve the problem.

We tested the proposed approach on a large-scale real-world network in Kyoto City with 1,189 links. A supply disruption scenario simulating the vehicle ban implemented during the Gion Matsuri (a cultural festival) was specifically designed for testing. The performance was assessed under three lane reversal control scenarios, i.e., 10, 15, and 20 controllable links. The experiment results suggested that the scenarios with 15 and 20 controlled links led to similar ultimate improvements in traffic resilience, but the 15-link scenario is more robust over algorithm iterations. This insight underscores the importance of striking a balance between the number of controllable links and model robustness. While this specific number is only applicable to the case study investigated in this study, we can extend these case-specific findings to a general discussion. Namely, insufficient controllable links (representing fewer control crews) cannot guarantee considerable mitigation in traffic resilience loss, while excessive controllable links would introduce too many degrees of freedom to the GPR model, which, however, can easily cause overfitting and result in a lack of robustness.

Further, we analyzed the evolution process of the surrogate model in detail from various perspectives. First, for a given time interval and a given control scenario, the surrogate model became more and more accurate in estimating simulation results, represented by a reduction in the variance of the distribution of their difference. Second, for a given control scenario, the variance of the difference between simulated and approximated resilience losses increased with time intervals, highlighting the long-lasting effect of control measures implemented in previous time intervals. In other words, temporally correlated traffic flows can transmit the effects of control measures across the network within a certain time period (multiple time intervals). Last but not least, we found that the distribution variance also increased with the number of controlled links. This finding is in line with the analysis mentioned above, i.e., excessive controllable links can cause overfitting and frustration in surrogate approximations.

Additionally, the structure of the proposed surrogate model was examined by a series of ablation experiments. Results confirmed the necessity of all elements contained in the surrogate model, including the GPR functional component, the physical component,

and the scaling factor for addressing systematic bias in the presented dynamic analytical network model. Moreover, the integration of a transition term in the analytical network model has also proven to be useful. Namely, the proposed dynamic network model is capable of capturing the temporal correlation of traffic states in different time intervals, thereby enhancing the approximation accuracy of the model.

Future research could extend this approach to incorporate budget constraints and integrate effective dimension reduction techniques to account for the long-lasting effects of recovery measures while preventing overfitting. Furthermore, exploring the integration of multiple practical recovery measures, such as traffic signal control, route guidance, and lane reversal, holds promise for providing comprehensive solutions in the disruption recovery phase of transportation systems.

Chapter 10

Conclusions, research findings and future work

Enhancing the resilience of transportation systems to various disruptions caused by natural disasters and human activities has emerged as a critical research topic within transportation engineering and urban studies. Given the pivotal role of transportation systems in facilitating the movement of people and goods across urban areas, this dissertation aims to investigate transportation system resilience through the lens of traffic flows. To this end, the concept of traffic resilience is employed and further refined. The dissertation, drawing insights from relevant literature, identifies three essential modeling components for comprehensive traffic resilience assessment frameworks: dynamic travel demand modeling, traffic resilience indicators, and traffic resilience optimization methodology. These components are interconnected, with the first two serving as prerequisites for the latter. Inspired by the conceptual relation diagram delineating the composition and structure of this framework, the dissertation contributes to the field by conducting studies that explore suitable data sources, appropriate modeling frameworks, efficient solution algorithms, and realistic scenario designs for each component. More specifically, the dissertation enriches the body of knowledge on the utilization of opportunistic data for dynamic demand estimation, introduces traffic resilience indicators based on MFDs, and proposes simulation-based approaches for traffic resilience optimization. Therefore, the dissertation can provide all stakeholders with valuable theoretical, methodological, and practical insights in traffic resilience evaluation and optimization.

10.1 Summary of main research findings

In the following, we present the main research findings extracted from our work specific to each component separately to offer a concise presentation.

10.1.1 Dynamic demand modeling using opportunistic data

- The dynamic OD estimator proposed in Chapter 4, driven by LBSN data and grounded on the scenario-based TSSP framework, establishes connections between inter-zone OD flows and zonal activity flows. Its notable performance in replicating inter-zone OD matrices underscores the feasibility of deriving OD flows from activity-level mobility flows and activity chain information.

- Employing GLS for model performance assessment, the proposed solution algorithm, which integrates GBD and MCMC sampling, demonstrates rapid convergence and yields a robust fit for check-in patterns, zonal production, zonal attraction, and OD flows.

- The exploration of the objective formulation space suggests that, due to the interdependence of zonal production, zonal attraction, and OD matrix, it is sufficient to involve two of them in the objective function, whereas the excluded term can serve as a validation metric for model outputs.

- Since the proposed LBSN-based OD estimator is simulation-free and independent of network structure, its computational cost is almost the bare minimum and only depends on the dimension of the OD matrix and the complexity of the graph models of traffic analysis zones. Therefore, it can effectively address the common indeterminateness issue faced by conventional traffic measurement-based OD estimators.

- Crowdsensed data harbor abundant real-time information regarding urban mobility patterns and are sensitive to system disruptions, holding substantial potential for basing point-level demand estimation modeling, especially during disruption periods when conventional data might be unavailable due to infrastructure damage.

- PT demand pattern changes caused by demand disruptions vary across different PT stations owing to their discrepancy in location, surrounding activity categories, local population, and daily demand patterns under normal conditions.

- Opportunistic data represent a valuable alternative for complementing, supplementing, or potentially replacing conventional survey-based and traffic measurement data in dynamic travel demand estimation up to a certain extent, contingent upon data quality and coverage levels.

10.1.2 MFD-based traffic resilience indicators

- Demand-side and supply-side disruptions manifest distinct impacts on transportation systems and influence traffic resilience differently. Supply-side disruptions will alter the shape of the MFD, often resulting in changes to traffic flow patterns, while demand-side disruptions, such as traffic congestion, do not affect the MFD shape but rather contribute to congestion occurrence.

- Traffic resilience indicators derived from MFD dynamics, as proposed in Chapter 6, are capable of integrating transportation network topology characteristics, traffic dynamics, and travel demand patterns. Compared to traditional topology-based and trip-based indicators, as well as prior MFD-based indicators, they exhibit superior reliability and robustness.

- Traffic resilience loss can be interpreted as the cumulative number of vehicles that should have finished their trips within the respective time intervals if the transportation system/network were operating optimally under normal conditions (without infrastructure malfunction).

- Grid-like networks demonstrate greater resilience to supply-side disruptions in comparison to networks with a central ring structure. This resilience is attributed to two factors: (i) Grid-like networks offer more route alternatives between OD pairs, facilitating trip rerouting post-disruption; (ii) The grid structure can ease vehicle detours.

- Topological attributes can, to a certain extent, explain traffic resilience, with variations across networks with different topology layouts. Among these attributes, network connectivity emerges as the most critical factor in ensuring transportation system resilience to supply disruptions.

10.1.3 Simulation-based traffic resilience optimization

- Surrogate models can integrate the strengths of traffic simulators in estimating time-dependent traffic dynamics and the superiority of mathematical optimization in computational efficiency.

- In the physics-informed surrogate model proposed in Chapter 8 for traffic resilience optimization, the physical component derived from a deterministic analytical network model provides valuable structural information on the relationship between traffic flow variables. Concurrently, the functional component is effective in modeling the difference between simulated and analytical outputs.

- For a specific recovery strategy scenario, with the number of simulation iterations increasing, the surrogate model gradually improves its approximation of the traffic simulator in terms of model outputs.

- Generally, the impact of control measures can last for a relatively longer period than the duration they are under implementation, representing the need for incorporating decision variables from previous time intervals in the fitting process of the functional component. This enables capturing the influence of control measures on traffic flow in subsequent time intervals.

- Lane reversal control is an effective and easy-to-implement recovery measure for improving urban traffic resilience in response to supply disruptions. However, it is important to strike a balance between the number of controlled links and model robustness; namely, insufficient controlled links would lead to a compromised reduction in traffic resilience loss, while excessive controlled links would easily cause overfitting and a loss of model robustness.

- Adding a traffic flow transition term to the analytical network model can capture more temporal correlation between traffic flows across different time intervals, thereby enhancing the approximation accuracy of the model.

10.2 Limitations and future works

While this dissertation develops various modeling frameworks and solution algorithms for dynamic demand estimation with opportunistic data, MFD-based traffic resilience indicators, and simulation-based traffic resilience optimization, and consolidates the methodological framework for traffic resilience evaluation and optimization, it is important to acknowledge that there are still some limitations in these studies that warrant further investigation in future research. Since the limitations of each study have been discussed in detail in their respective chapters, this section aims to highlight additional aspects that have not been fully addressed in the dissertation but are crucial for enhancing our understanding of urban traffic resilience and its practical applications in real operational, tactical, and strategic scenarios. In the following subsections, we identify and discuss the limitations of each component and suggest potential future research directions to address these limitations.

10.2.1 Dynamic demand modeling using opportunistic data

- The inevitable sampling bias exiting in LBSN data due to the higher engagement of young people in social media compared to other age groups may introduce inaccuracies and render the estimation results from the proposed LBSN-based OD estimator unrepresentative. Therefore, a promising avenue for future research is to devise a calibration method for the scaling factor matrix to ensure that the scaled estimation results remain accurate, thereby enhancing the practical value of the LBSN-based OD estimator. Another way around to address this issue is to integrate the proposed model with an LBSN simulation model. We believe that an LBSN simulation model calibrated with the observed LBSN data has eliminated such sampling bias inherent in LBSN data, and therefore can produce reliable simulation data closing to realistic situations. This challenge is not unique to LBSN data and is also pertinent to other similar opportunistic data, such as crowdsensed data. Hence, these two solutions can be extended to address similar issues encountered with other types of opportunistic data.

- Another apparent limitation of opportunistic data in dynamic demand estimation is the absence of ground truth values, imposing a critical challenge in model validation. To this end, these models must be linked to another regression model describing the relationship between opportunistic data observations and other data sources with ground truth values. For instance, Vongvanich et al. (2023) developed a stepwise multiple linear regression model to estimate the correlation between the true PT demand and the busyness rates of the POIs in the station vicinity.

- While we have presented the potential of opportunistic data for dynamic demand estimation, combining these data with conventional sources of data like traffic measurements through data fusion techniques could further strengthen the reliability and robustness of dynamic demand estimation frameworks. Such integrated frameworks are versatile and applicable across various scenarios, as it is nearly impossible that all data sources would be unavailable simultaneously.

- Due to the lack of relevant data, we could not examine the performance of the proposed models in the context of supply disruptions. As a result, it would also be interesting to collect some opportunistic data during transportation supply disruptions and test our models in such scenarios.

10.2.2 MFD-based traffic resilience indicators

- The interplay between a supply-side disruption and the induced emergency demand it creates often leads to mixed disruption scenarios. While this dissertation has provided a method to assess traffic resilience loss under such scenarios, no experiments have been conducted to analyze their resilience loss patterns and compare them with scenarios involving a single type of disruption. Therefore, an interesting direction for future studies would be to design plausible combinations of disruptions and explore the unique characteristics of traffic resilience under these mixed disruption scenarios.

- In the dissertation, our analyses of traffic resilience are based on traffic simulation results, which can provide useful insights into the practical application of the concept of traffic resilience. However, it is equally crucial to deduce some fundamental physical properties of traffic resilience for transportation systems with different MFD curves, such as parabola-shaped MFDs versus triangle-shaped MFDs, through analytical modeling, as demonstrated by Gao et al. (2022) in their phase diagram analysis. To achieve this, developing smooth approximations to the proposed non-smooth indicators would be beneficial.

- The regression analysis performed in Chapter 7 is limited to two specific types of network topology due to data constraints. Future research endeavors can encompass more categories of network structures present in various cities around the world to explore the potential existence of a general principle relationship between traffic resilience and network topology. To accomplish this, a key challenge is the creation of reliable synthetic network generation models, with the difficulty lies in creating reasonable OD demand matrices for comparative analysis. The prospect of such a model is promising, given its helpfulness in the design of resilient transportation networks.

- In Chapter 5, we developed a framework to estimate PT demand pattern changes contributed by demand disruptions. However, private cars and PT are not distinguished in the development of MFD-based indicators for simplicity. While this

simplification was practical, enhancing the definition of urban traffic resilience and its indicators would benefit from considering the context of multimodal transportation systems. For instance, adopting the 3D-MFD proposed in Loder et al. (2017) could replace the MFD specifically for private cars, thus providing a more nuanced understanding of system dynamics in mixed traffic environments.

10.2.3 Simulation-based traffic resilience optimization

- Although the analytical network model developed in Chapter 8 is sufficiently efficient compared with traffic simulations, more efforts can still be directed towards simplifying its model structure and improving its implementation in programming languages to reduce its computational cost even further.
- For simplicity, in Chapter 9, we formulated a recovery measure optimization problem with a personnel constraint, namely, a limitation on the number of controllable links. However, acknowledging that real-world implementations may entail additional constraints such as financial limitations, future research could extend the proposed model to incorporate these constraints. Given the increased complexity, novel efficient algorithms would be necessary.
- In the experiments detailed in Chapter 9, we discarded control variables from previous time intervals when fitting the functional component for traffic resilience loss within the current interval, though recognizing the long-lasting effects of these variables. This treatment prevents overfitting, which on the other hand, sacrifices some accuracy. Thus, applying appropriate dimension reduction techniques to the set of control variables for all time intervals might be useful to address this problem.
- Another interesting direction for future research is to explore the efficacy of combinations of different types of recovery measures in addition to the lane reversal control considered in our study. Examples include traffic signal control, route guidance, and additional lane reversal strategies.
- Beyond localized link closures associated with supply disruptions, it would be valuable to evaluate our approaches in other disruption scenarios such as network-wide flooding. The broader examination would enhance the robustness and applicability of the methods presented herein.
- In addition to recovery measures, which have an effect on the recovery phase of the disruption, preparedness measures are also critical to the capability of transportation systems to withstand various disruptions. One popular research topic in preparedness optimization lies in Road Network Design Problems (RNDPs). In contrast to traditional methods relying on bi-level optimization frameworks coupled with analytical UE constraints, future research can be conducted to explore the application of traffic simulation in addressing this problem at strategic, tactical, and operational decision levels (Farahani et al., 2013).

Bibliography

Ahmed, M. A., Sadri, A. M., Mehrabi, A., and Azizinamini, A. (2022). Identifying topological credentials of physical infrastructure components to enhance transportation network resilience: case of Florida bridges. *Journal of Transportation Engineering, Part A: Systems*, 148(9):04022055.

Ahmed, S. and Dey, K. (2020). Resilience modeling concepts in transportation systems: A comprehensive review based on mode, and modeling techniques. *Journal of Infrastructure Preservation and Resilience*, 1(1):8.

Alkhaleel, B. A., Liao, H., and Sullivan, K. M. (2022). Risk and resilience-based optimal post-disruption restoration for critical infrastructures under uncertainty. *European Journal of Operational Research*, 296(1):174–202.

Amghar, R., Jaber, S., Moghaddam, S. H. M., Bhouri, N., and Ameli, M. (2024). Resilience as a service for transportation networks: Definition and basic concepts. *Transportation Research Record*, 2678(1):177–189.

Amini, K. and Padgett, J. E. (2023). Probabilistic risk assessment of hurricane-induced debris impacts on coastal transportation infrastructure. *Reliability Engineering & System Safety*, 240:109579.

Amini, S., Tilg, G., and Busch, F. (2018). Evaluating the impact of real-time traffic control measures on the resilience of urban road networks. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 519–524.

Anderson, M., Kiddle, D., and Logan, T. (2022). The underestimated role of the transportation network: Improving disaster & community resilience. *Transportation Research Part D: Transport and Environment*, 106:103218.

Antoniou, C., Azevedo, C. L., Lu, L., Pereira, F., and Ben-Akiva, M. (2015). W-SPSA in practice: Approximation of weight matrices and calibration of traffic simulation models. *Transportation Research Part C: Emerging Technologies*, 59:129–146.

Antoniou, C., Barceló, J., Breen, M., Bullejos, M., Casas, J., Cipriani, E., Ciuffo, B., Djukic, T., Hoogendoorn, S., Marzano, V., Montero, L., Nigro, M., Perarnau, J., Punzo, V., Toledo, T., and van Lint, H. (2016). Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation. *Transportation Research Part C: Emerging Technologies*, 66:79–98.

Aparicio, J. T., Arsenio, E., and Henriques, R. (2022). Assessing robustness in multimodal transportation systems: a case study in Lisbon. *European Transport Research Review*, 14(1):1–18.

Arango, E., Nogal, M., Yang, M., Sousa, H. S., Stewart, M. G., and Matos, J. C. (2023). Dynamic thresholds for the resilience assessment of road traffic networks to wildfires. *Reliability Engineering & System Safety*, page 109407.

Au, S.-K. and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277.

Aydin, N. Y., Duzgun, H. S., Heinimann, H. R., Wenzel, F., and Gnyawali, K. R. (2018). Framework for improving the resilience and recovery of transportation networks under geohazard risks. *International Journal of Disaster Risk Reduction*, 31:832–843.

Balakrishna, R., Antoniou, C., Ben-Akiva, M., Koutsopoulos, H. N., and Wen, Y. (2007). Calibration of microscopic traffic simulation models: Methods and application. *Transportation Research Record*, 1999(1):198–207.

Balal, E., Valdez, G., Miramontes, J., and Cheu, R. L. (2019). Comparative evaluation of measures for urban highway network resilience due to traffic incidents. *International Journal of Transportation Science and Technology*, 8(3):304–317.

Bauer, D., Richter, G., Asamer, J., Heilmann, B., Lenz, G., and Kölbl, R. (2017). Quasi-dynamic estimation of OD flows from traffic counts without prior OD matrix. *IEEE Transactions on Intelligent Transportation Systems*, 19(6):2025–2034.

Beglerovic, H., Stolz, M., and Horn, M. (2017). Testing of autonomous vehicles using surrogate models and stochastic optimization. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6. IEEE.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

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

Bojer, C. S. and Meldgaard, J. P. (2021). Kaggle forecasting competitions: An overlooked learning opportunity. *International Journal of Forecasting*, 37(2):587–603.

Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A., and Von Winterfeldt, D. (2003). A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra*, 19(4):733–752.

Bucar, R. C. and Hayeri, Y. M. (2020). Quantitative assessment of the impacts of disruptive precipitation on surface transportation. *Reliability Engineering & System Safety*, 203:107105.

Byun, J.-E. and D'Ayala, D. (2022). Urban seismic resilience mapping: a transportation network in Istanbul, Turkey. *Scientific Reports*, 12(1):8188.

Cai, Z., Wang, D., Chen, X. M., et al. (2017). A novel trip coverage index for transit accessibility assessment using mobile phone data. *Journal of Advanced Transportation*, 2017.

Cantelmo, G., Qurashi, M., Prakash, A. A., Antoniou, C., and Viti, F. (2020). Incorporating trip chaining within online demand estimation. *Transportation Research Part B: Methodological*, 132:171–187.

Cantelmo, G., Viti, F., Tampère, C. M., Cipriani, E., and Nigro, M. (2014). Two-step approach for correction of seed matrix in dynamic demand estimation. *Transportation Research Record*, 2466(1):125–133.

Capponi, A., Vitello, P., Fiandrino, C., Cantelmo, G., Kliazovich, D., Sorger, U., and Bouvry, P. (2019). Crowdsensed data learning-driven prediction of local businesses attractiveness in smart cities. In *2019 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6. IEEE.

Cascetta, E., Papola, A., Marzano, V., Simonelli, F., and Vitiello, I. (2013). Quasi-dynamic estimation of o–d flows from traffic counts: Formulation, statistical validation and performance analysis on real data. *Transportation Research Part B: Methodological*, 55:171–187.

Cascetta, E. and Postorino, M. N. (2001). Fixed point approaches to the estimation of O/D matrices using traffic counts on congested networks. *Transportation Science*, 35(2):134–147.

Cebelak, M. K. (2013). Location-based social networking data: Doubly-constrained gravity model origin-destination estimation of the urban travel demand for Austin, TX. Master's thesis, The University of Texas at Austin, United States.

Chaniotakis, E., Antoniou, C., and Pereira, F. C. (2017). Enhancing resilience to disasters using social media. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 699–703.

Chen, C., Huang, Y., Lam, W., Pan, T., Hsu, S., Sumalee, A., and Zhong, R. (2022). Data efficient reinforcement learning and adaptive optimal perimeter control of network traffic dynamics. *Transportation Research Part C: Emerging Technologies*, 142:103759.

Chen, C., Wang, S., Zhang, J., and Gu, X. (2023). Modeling the vulnerability and resilience of interdependent transportation networks under multiple disruptions. *Journal of Infrastructure Systems*, 29(1):04022043.

Chen, L. and Miller-Hooks, E. (2012). Resilience: An indicator of recovery capability in intermodal freight transport. *Transportation Science*, 46(1):109–123.

Chen, X. M., Zhang, L., He, X., Xiong, C., and Li, Z. (2014). Surrogate-based optimization of expensive-to-evaluate objective for optimal highway toll charges in transportation network. *Computer-Aided Civil and Infrastructure Engineering*, 29(5):359–381.

Chen, X. M., Zhu, Z., He, X., and Zhang, L. (2015). Surrogate-based optimization for solving a mixed integer network design problem. *Transportation Research Record*, 2497(1):124–136.

Cho, Y.-S., Steeg, G. V., and Galstyan, A. (2014). Where and why users "check in". In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 269–275.

Chong, L. and Osorio, C. (2018). A simulation-based optimization algorithm for dynamic large-scale urban transportation problems. *Transportation Science*, 52(3):637–656.

Cools, M., Fabbro, Y., and Bellemans, T. (2016). Free public transport: A socio-cognitive analysis. *Transportation Research Part A: Policy and Practice*, 86:96–107.

Dadashzadeh, N., Elvarsson, A. B., Morshed, G., Agriesti, S. A., Antoniou, C., Roncoli, C., and Thomopoulos, N. (2021). Autonomous and connected transport scenarios evaluation based on simulation analysis: Wg5: Thematic report. Technical report, European Cooperation in Science & Technology (COST).

Diab, E. and Shalaby, A. (2020). Metro transit system resilience: Understanding the impacts of outdoor tracks and weather conditions on metro system interruptions. *International Journal of Sustainable Transportation*, 14(9):657–670.

Dingil, A. E., Rupi, F., and Stasiskiene, Z. (2019). A macroscopic analysis of transport networks: The influence of network design on urban transportation performance. *International Journal of Transport Development and Integration*, 3(4):331–343.

Djukic, T., Van Lint, J., and Hoogendoorn, S. (2012). Application of principal component analysis to predict dynamic origin–destination matrices. *Transportation Research Record*, 2283(1):81–89.

Dong, S., Gao, X., Mostafavi, A., Gao, J., and Gangwal, U. (2023). Characterizing resilience of flood-disrupted dynamic transportation network through the lens of link reliability and stability. *Reliability Engineering & System Safety*, 232:109071.

D'Silva, K., Noulas, A., Musolesi, M., Mascolo, C., and Sklar, M. (2018). Predicting the temporal activity patterns of new venues. *EPJ Data Science*, 7(1):13.

Fan, Z., Zhang, F., and Loo, B. P. Y. (2022). Rhythm of transit stations - Uncovering the activity-travel dynamics of transit-oriented development in the U.S. *IEEE Transactions on Intelligent Transportation Systems*, 23(8):12503–12517.

Fang, C., Chu, Y., Fu, H., and Fang, Y. (2022). On the resilience assessment of complementary transportation networks under natural hazards. *Transportation Research Part D: Transport and Environment*, 109:103331.

Farahani, R. Z., Miandoabchi, E., Szeto, W., and Rashidi, H. (2013). A review of urban transportation network design problems. *European Journal of Operational Research*, 229(2):281–302.

Faturechi, R. and Miller-Hooks, E. (2014). Travel time resilience of roadway networks under disaster. *Transportation Research Part B: Methodological*, 70:47–64.

FHWA (2015). Transportation system resilience to extreme weather and climate change. `https://ops.fhwa.dot.gov/publications/fhwahop15025/fhwahop15025.pdf` (Last accessed on 2023-06-20).

Forrester, A., Sobester, A., and Keane, A. (2008). *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons.

Frederix, R., Viti, F., Corthout, R., and Tampère, C. M. (2011). New gradient approximation method for dynamic origin–destination matrix estimation on congested networks. *Transportation Research Record*, 2263(1):19–25.

Ganin, A. A., Mersky, A. C., Jin, A. S., Kitsak, M., Keisler, J. M., and Linkov, I. (2019). Resilience in intelligent transportation systems (ITS). *Transportation Research Part C: Emerging Technologies*, 100:318–329.

Gao, J., Barzel, B., and Barabási, A.-L. (2016). Universal resilience patterns in complex networks. *Nature*, 530(7590):307–312.

Gao, S., Li, D., Zheng, N., Hu, R., and She, Z. (2022). Resilient perimeter control for hypercongested two-region networks with MFD dynamics. *Transportation Research Part B: Methodological*, 156:50–75.

Ge, Y., Du, L., and Ye, H. (2019). Co-optimization approach to post-storm recovery for interdependent power and transportation systems. *Journal of Modern Power Systems and Clean Energy*, 7(4):688–695.

Geoffrion, A. M. (1972). Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260.

Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9):759–770.

Gonçalves, L. and Ribeiro, P. (2020). Resilience of urban transportation systems. Concept, characteristics, and methods. *Journal of Transport Geography*, 85:102727.

Google (2023). Popular times, wait times, and visit duration. Accessed on 07.04.2023.

Google Maps (2023). Google maps. Accessed on 07.04.2023.

Guidotti, R., Gardoni, P., and Chen, Y. (2017). Network reliability analysis with link and nodal weights and auxiliary nodes. *Structural Safety*, 65:12–26.

Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). Exploring network structure, dynamics, and function using networkx. In Varoquaux, G., Vaught, T., and Millman, J., editors, *Proceedings of the 7th Python in Science Conference*, pages 11 – 15, Pasadena, CA USA.

Hao, Y., Jia, L., Zio, E., Wang, Y., Small, M., and Li, M. (2023). Improving resilience of high-speed train by optimizing repair strategies. *Reliability Engineering & System Safety*, 237:109381.

Harrison, G., Grant-Muller, S. M., and Hodgson, F. C. (2020). New and emerging data forms in transportation planning and policy: Opportunities and challenges for "track and trace" data. *Transportation Research Part C: Emerging Technologies*, 117:102672.

HDX (2023). Germany: High resolution population density maps + demographic estimates. Accessed on 20.04.2023.

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4(1):1–23.

Hoogendoorn, S. P., Knoop, V. L., van Lint, H., and Vu, H. L. (2015). Applications of the generalized macroscopic fundamental diagram. In Chraibi, M., Boltes, M., Schadschneider, A., and Seyfried, A., editors, *Traffic and Granular Flow'13*, pages 577–583. Springer International Publishing.

Hu, N. W. and Jin, P. J. (2015). Dynamic trip attraction estimation with location based social network data balancing between time of day variations and zonal differences. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(4):193.

Hu, W. and Jin, P. J. (2017). An adaptive hawkes process formulation for estimating time-of-day zonal trip arrivals with location-based social networking check-in data. *Transportation Research Part C: Emerging Technologies*, 79:136–155.

Hu, W., Yao, Z., Yang, S., Chen, S., and Jin, P. J. (2019). Discovering urban travel demands through dynamic zone correlation in location-based social networks. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018*, pages 88–104. Springer.

Huang, J. and Levinson, D. M. (2015). Circuity in urban transit networks. *Journal of Transport Geography*, 48:145–153.

Huang, Y., Xiong, J., Sumalee, A., Zheng, N., Lam, W., He, Z., and Zhong, R. (2020). A dynamic user equilibrium model for multi-region macroscopic fundamental diagram systems with time-varying delays. *Transportation Research Part B: Methodological*, 131:1–25.

Huo, J., Liu, C., Chen, J., Meng, Q., Wang, J., and Liu, Z. (2023). Simulation-based dynamic origin-destination matrix estimation on freeways: A Bayesian optimization approach. *Transportation Research Part E: Logistics and Transportation Review*, 173:103108.

Jana, D., Malama, S., Narasimhan, S., and Taciroglu, E. (2023). Edge-based graph neural network for ranking critical road segments in a network. *Plos One*, 18(12):e0296045.

Jin, P. J., Cebelak, M., Yang, F., Zhang, J., Walton, C. M., and Ran, B. (2014). Location-based social networking data: exploration into use of doubly constrained gravity model for origin–destination estimation. *Transportation Research Record*, 2430(1):72–82.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3):187–200.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems 30*, pages 3146–3154. Curran Associates, Inc.

Kheiri, A., Karimipour, F., and Forghani, M. (2015). Intra-urban movement flow estimation using location based social networking data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(1):781.

Kim, S. and Yeo, H. (2017). Evaluating link criticality of road network based on the concept of macroscopic fundamental diagram. *Transportmetrica A: Transport Science*, 13(2):162–193.

Kurth, M., Kozlowski, W., Ganin, A., Mersky, A., Leung, B., Dykes, J., Kitsak, M., and Linkov, I. (2020). Lack of resilience in transportation networks: Economic implications. *Transportation Research Part D: Transport and Environment*, 86:102419.

Levinson, D. (2012). Network structure and city size. *PloS One*, 7(1):e29721.

Li, Z., Jin, C., Hu, P., and Wang, C. (2019). Resilience-based transportation network recovery strategy during emergency recovery phase under uncertainty. *Reliability Engineering & System Safety*, 188:503–514.

Li, Z., Tang, W., Lian, X., Chen, X., Zhang, W., and Qian, T. (2022). A resilience-oriented two-stage recovery method for power distribution system considering transportation network. *International Journal of Electrical Power & Energy Systems*, 135:107497.

Liao, T.-Y., Hu, T.-Y., and Ko, Y.-N. (2018). A resilience optimization model for transportation networks under disasters. *Natural Hazards*, 93(1):469–489.

Liu, K., Zhai, C., and Dong, Y. (2021). Optimal restoration schedules of transportation network considering resilience. *Structure and Infrastructure Engineering*, 17(8):1141–1154.

Lo, H. K. and Tung, Y.-K. (2003). Network with degradable links: Capacity analysis and design. *Transportation Research Part B: Methodological*, 37(4):345–363.

Loder, A., Ambühl, L., Menendez, M., and Axhausen, K. W. (2017). Empirics of multi-modal traffic networks–using the 3d macroscopic fundamental diagram. *Transportation Research Part C: Emerging Technologies*, 82:88–101.

Loder, A., Cantner, F., Adenaw, L., and Bogenberger, K. (2023). The 9 EUR-ticket -A nationwide experiment: Germany -first findings. In *102nd Annual Meeting of the Transportation Research Board (TRB 2023)*.

López, E. and Monzón, A. (2010). Integration of sustainability issues in strategic transportation planning: a multi-criteria model for the assessment of transport infrastructure plans. *Computer-Aided Civil and Infrastructure Engineering*, 25(6):440–451.

Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., and Wießner, E. (2018). Microscopic traffic simulation using SUMO. In *2018 21st international conference on intelligent transportation systems (ITSC)*, pages 2575–2582. IEEE.

Lu, Q.-L., Mahajan, V., Lyu, C., and Antoniou, C. (2023a). Analyzing crowding in public transport during 9-EUR ticket using crowdsensing data. In *25th Euro Working Group on Transportation Meeting (EWGT 2023)*.

Lu, Q.-L., Mahajan, V., Lyu, C., and Antoniou, C. (2024a). Analyzing the impact of fare-free public transport policies on crowding patterns at stations using crowdsensing data. *Transportation Research Part A: Policy and Practice*, 179:103944.

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2022). A stochastic programming method for od estimation using lbsn check-in data. In *4th Symposium on Management of Future Motorway and Urban Traffic Systems (MFTS 2022)*.

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2023b). Simulation-based policy analysis: The case of urban speed limits. *Transportation Research Part A: Policy and Practice*, 175:103754.

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2023c). A two-stage stochastic programming approach for dynamic OD estimation. In *102nd Annual Meeting of the Transportation Research Board (TRB 2023)*.

Lu, Q.-L., Qurashi, M., and Antoniou, C. (2024b). A two-stage stochastic programming approach for dynamic OD estimation using LBSN data. *Transportation Research Part C: Emerging Technologies*, 158:104460.

Lu, Q.-L., Qurashi, M., Varesanovic, D., Sodnik, J., and Antoniou, C. (2021). Exploring the influence of automated driving styles on network efficiency. *Transportation Research Procedia*, 52:380–387.

Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2023d). An MFD-based optimization approach to improve transportation system resilience under infrastructure disruption. In *25th Euro Working Group on Transportation Meeting (EWGT 2023)*.

Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2023e). Surrogate modeling for recovery measure optimization to improve traffic resilience. In *9th International Symposium on Transport Network Resilience (INSTR 2023)*.

Lu, Q.-L., Sun, W., Dai, J., Schmöcker, J.-D., and Antoniou, C. (2024c). Traffic resilience quantification based on macroscopic fundamental diagrams and analysis using topological attributes. *Reliability Engineering & System Safety*, 247:110095.

Lu, Q.-L., Sun, W., Lyu, C., Schmöcker, J.-D., and Antoniou, C. (2024d). Post-disruption lane reversal optimization with surrogate model to improve urban traffic resilience. *(Under Review)*.

MacKenzie, D. and Cho, H. (2020). Travel demand and emissions from driving dogs to dog parks. *Transportation Research Record*, 2674(6):291–296.

Mahajan, V. (2023). *Data openness and efficient methods to address data insufficiency in mobility analyses and simulation calibration*. PhD thesis, Technische Universität München.

Mahajan, V., Cantelmo, G., and Antoniou, C. (2021). Explaining demand patterns during COVID-19 using opportunistic data: A case study of the city of Munich. *European Transport Research Review*, 13(1):26.

Mahajan, V., Cantelmo, G., Rothfeld, R., and Antoniou, C. (2023). Predicting network flows from speeds using open data and transfer learning. *IET Intelligent Transport Systems*, 17(4):804–824.

McNally, M. G. (2007). The four-step model. In *Handbook of transport modelling*. Emerald Group Publishing Limited.

Miller-Hooks, E., Zhang, X., and Faturechi, R. (2012). Measuring and maximizing resilience of freight transportation networks. *Computers & Operations Research*, 39(7):1633–1643.

Mirjalili, R., Barati, H., and Yazici, A. (2023). Resilience analysis of New York City transportation network after snow storms. *Transportation Research Record*, 2677(1):694–707.

Mnih, A. and Salakhutdinov, R. R. (2007). Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*, volume 20. Curran Associates, Inc.

Möhring, M., Keller, B., Schmidt, R., and Dacko, S. (2021). Google Popular Times: Towards a better understanding of tourist customer patronage behavior. *Tourism Review*, 76(3):533–569.

Nan, C. and Sansavini, G. (2017). A quantitative method for assessing resilience of interdependent infrastructures. *Reliability Engineering & System Safety*, 157:35–53.

Niu, H. and Silva, E. A. (2020). Crowdsourced data mining for urban activity: Review of data sources, applications, and methods. *Journal of Urban Planning and Development*, 146(2):04020007.

NOAA, U. (2024). Billion-dollar weather and climate disasters. `https://www.ncei.noaa.gov/access/billions/events/US/2023` (Last accessed on 2024-02-15).

Osorio, C. (2019a). Dynamic origin-destination matrix calibration for large-scale network simulators. *Transportation Research Part C: Emerging Technologies*, 98:186–206.

Osorio, C. (2019b). High-dimensional offline origin-destination (OD) demand calibration for stochastic traffic simulators of large-scale road networks. *Transportation Research Part B: Methodological*, 124:18–43.

Osorio, C. and Bierlaire, M. (2009). An analytic finite capacity queueing network model capturing the propagation of congestion and blocking. *European Journal of Operational Research*, 196(3):996–1007.

Osorio, C. and Bierlaire, M. (2013). A simulation-based optimization framework for urban transportation problems. *Operations Research*, 61(6):1333–1345.

Osorio, C. and Chong, L. (2015). A computationally efficient simulation-based optimization algorithm for large-scale urban transportation problems. *Transportation Science*, 49(3):623–636.

Pan, S., Yan, H., He, J., and He, Z. (2021). Vulnerability and resilience of transportation systems: A recent literature review. *Physica A: Statistical Mechanics and its Applications*, 581(June).

Parthasarathi, P. (2014). Network structure and metropolitan mobility. *Journal of Transport and Land Use*, 7(2):153–170.

Pei, S., Zhai, C., and Hu, J. (2024). Surrogate model-assisted seismic resilience assessment of the interdependent transportation and healthcare system considering a two-stage recovery strategy. *Reliability Engineering & System Safety*, 244:109941.

Pender, B., Currie, G., Delbosc, A., and Shiwakoti, N. (2014). Social media use during unplanned transit network disruptions: A review of literature. *Transport Reviews*, 34(4):501–521.

Pinelli, F., Nair, R., Calabrese, F., Berlingerio, M., Di Lorenzo, G., and Sbodio, M. L. (2016). Data-driven transit network design from mobile phone trajectories. *IEEE Transactions on Intelligent Transportation Systems*, 17(6):1724–1733.

Qurashi, M., Lu, Q.-L., Cantelmo, G., and Antoniou, C. (2022). Dynamic demand estimation on large scale networks using principal component analysis: The case of non-existent or irrelevant historical estimates. *Transportation Research Part C: Emerging Technologies*, 136:103504.

Qurashi, M., Ma, T., Chaniotakis, E., and Antoniou, C. (2019). PC–SPSA: employing dimensionality reduction to limit SPSA search noise in DTA model calibration. *IEEE Transactions on Intelligent Transportation Systems*, 21(4):1635–1645.

Rasmussen, C. E. and Williams, C. K. I. (2008). *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, 3. print edition.

Reynolds, D. A. et al. (2009). Gaussian mixture models. *Encyclopedia of biometrics*, 741(659-663).

Rizwan, M., Wan, W., and Gwiazdzinski, L. (2020). Visualization, spatiotemporal patterns, and directional analysis of urban activities using geolocation data extracted from LBSN. *ISPRS International Journal of Geo-Information*, 9(2):137.

Roy, K. C., Cebrian, M., and Hasan, S. (2019). Quantifying human mobility resilience to extreme events using geo-located social media data. *EPJ Data Science*, 8(1):1–15.

Safitri, N. D. and Chikaraishi, M. (2022). Impact of transport network disruption on travel demand: A case study of the July 2018 heavy rain disaster in Japan. *Asian Transport Studies*, 8:100057.

Santiago-Iglesias, E., Carpio-Pinedo, J., Sun, W., and García-Palomares, J. C. (2023). Frozen city: Analysing the disruption and resilience of urban activities during a heavy snowfall event using Google Popular Times. *Urban Climate*, 51:101644.

Serdar, M. Z., Koç, M., and Al-Ghamdi, S. G. (2022). Urban transportation networks resilience: indicators, disturbances, and assessment methods. *Sustainable Cities and Society*, 76:103452.

Shafiei, S., Saberi, M., Zockaie, A., and Sarvi, M. (2017). Sensitivity-based linear approximation method to estimate time-dependent origin–destination demand in congested networks. *Transportation Research Record*, 2669(1):72–79.

Shekar, V., Fiondella, L., Chatterjee, S., and Halappanavar, M. (2017). Quantitative assessment of transportation network vulnerability with dynamic traffic simulation methods. In *2017 IEEE International Symposium on Technologies for Homeland Security (HST)*, pages 1–7. IEEE.

Shwartz-Ziv, R. and Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81:84–90.

Silva, T. H., Viana, A. C., Benevenuto, F., Villas, L., Salles, J., Loureiro, A., and Quercia, D. (2019). Urban computing leveraging location-based social network data: A survey. *ACM Computing Surveys (CSUR)*, 52(1):1–39.

Somy, S., Shafaei, R., and Ramezanian, R. (2022). Resilience-based mathematical model to restore disrupted road-bridge transportation networks. *Structure and Infrastructure Engineering*, 18(9):1334–1349.

Spall, J. C. (1998). Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3):817–823.

Su, Z., Chow, A. H., Zheng, N., Huang, Y., Liang, E., and Zhong, R. (2020). Neuro-dynamic programming for optimal control of macroscopic fundamental diagram systems. *Transportation Research Part C: Emerging Technologies*, 116:102628.

Taghizadeh, M., Mahsuli, M., and Poorzahedy, H. (2023). Probabilistic framework for evaluating the seismic resilience of transportation systems during emergency medical response. *Reliability Engineering & System Safety*, 236:109255.

Thøgersen, J. (2009). Promoting public transport as a subscription service: Effects of a free month travel card. *Transport Policy*, 16(6):335–343.

Timokhin, S., Sadrani, M., and Antoniou, C. (2020). Predicting venue popularity using crowdsourced and passive sensor data. *Smart Cities*, 3(3):818–841.

Toledo, T. and Kolechkina, T. (2012). Estimation of dynamic origin–destination matrices using linear assignment matrix approximations. *IEEE Transactions on Intelligent Transportation Systems*, 14(2):618–626.

Tse, R., Zhang, L. F., Lei, P., and Pau, G. (2018). Social network based crowd sensing for intelligent transportation and climate applications. *Mobile Networks and Applications*, 23(1):177–183.

Tympakianaki, A., Koutsopoulos, H. N., and Jenelius, E. (2015). c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin–destination matrix estimation. *Transportation Research Part C: Emerging Technologies*, 55:231–245.

Virtucio, M. B., Cetiner, B., Zhao, B., Soga, K., and Taciroglu, E. (2024). A granular framework for modeling the capacity loss and recovery of regional transportation networks under seismic hazards: A case study on the port of los angeles. *International Journal of Disaster Risk Reduction*, 100:104164.

Vitello, P., Fiandrino, C., Connors, R. D., and Viti, F. (2023). TransitCrowd: Estimating subway stations demand with mobile crowdsensing data. preprint, In Review.

Vongvanich, T., Sun, W., and Schmöcker, J.-D. (2023). Explaining and predicting station demand patterns using Google Popular Times data. *Data Science for Transportation*, 5(2):10.

Wan, C., Yang, Z., Zhang, D., Yan, X., and Fan, S. (2018). Resilience in transportation systems: A systematic review and future directions. *Transport Reviews*, 38(4):479–498.

Wang, N., Wu, M., and Yuen, K. F. (2023). A novel method to assess urban multimodal transportation system resilience considering passenger demand and infrastructure supply. *Reliability Engineering & System Safety*, 238:109478.

Wang, Y. and Wang, J. (2020). Measuring and maximizing resilience of transportation systems for emergency evacuation. *IEEE Transactions on Engineering Management*, 67(3):603–613.

Wassmer, J., Merz, B., and Marwan, N. (2024). Resilience of transportation infrastructure networks to road failures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34(1).

Wei, D., Rose, A., Koc, E., Chen, Z., and Soibelman, L. (2022). Socioeconomic impacts of resilience to seaport and highway transportation network disruption. *Transportation Research Part D: Transport and Environment*, 106:103236.

Wu, Y. and Chen, S. (2023). Traffic resilience modeling for post-earthquake emergency medical response and planning considering disrupted infrastructure and dislocated residents. *International Journal of Disaster Risk Reduction*, 93:103754.

Xiao, Y., Zhao, X., Wu, Y., Chen, Z., Gong, H., Zhu, L., and Liu, Y. (2022). Seismic resilience assessment of urban interdependent lifeline networks. *Reliability Engineering & System Safety*, 218:108164.

Xiong, X., Ozbay, K., Jin, L., and Feng, C. (2020). Dynamic origin–destination matrix prediction with line graph neural networks and Kalman filter. *Transportation Research Record*, 2674(8):491–503.

Xu, J., Nelson, B. L., and Hong, L. J. (2013). An adaptive hyperbox algorithm for high-dimensional discrete optimization via simulation problems. *INFORMS Journal on Computing*, 25(1):133–146.

Xu, M., Ouyang, M., Hong, L., Mao, Z., and Xu, X. (2022). Resilience-driven repair sequencing decision under uncertainty for critical infrastructure systems. *Reliability Engineering & System Safety*, 221:108378.

Xu, Z. and Chopra, S. S. (2022). Network-based assessment of metro infrastructure with a spatial–temporal resilience cycle framework. *Reliability Engineering & System Safety*, 223:108434.

Yang, D., Zhang, D., Zheng, V. W., and Yu, Z. (2014a). Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(1):129–142.

Yang, F., Jin, P. J., Cheng, Y., Zhang, J., and Ran, B. (2015). Origin-destination estimation for non-commuting trips using location-based social networking data. *International Journal of Sustainable Transportation*, 9(8):551–564.

Yang, F., Jin, P. J., Wan, X., Li, R., and Ran, B. (2014b). Dynamic origin-destination travel demand estimation using location based social networking data. In *Transportation Research Board 93th Annual Meeting*.

Yang, Z., Barroca, B., Weppe, A., Bony-Dandrieux, A., Laffréchine, K., Daclin, N., November, V., Omrane, K., Kamissoko, D., Benaben, F., et al. (2023). Indicator-based resilience assessment for critical infrastructures–a review. *Safety Science*, 160:106049.

Yao, K. and Chen, S. (2023). Percolation-based resilience modeling and active intervention of disrupted urban traffic network during a snowstorm. *Journal of Transportation Engineering, Part A: Systems*, 149(5):04023027.

Yin, J., Ren, X., Liu, R., Tang, T., and Su, S. (2022). Quantitative analysis for resilience-based urban rail systems: A hybrid knowledge-based and data-driven approach. *Reliability Engineering & System Safety*, 219:108183.

Yin, K., Wu, J., Wang, W., Lee, D.-H., and Wei, Y. (2023). An integrated resilience assessment model of urban transportation network: A case study of 40 cities in China. *Transportation Research Part A: Policy and Practice*, 173:103687.

Yoon, S., Kim, J., Kim, M., Tak, H.-Y., and Lee, Y.-J. (2020). Accelerated system-level seismic risk assessment of bridge transportation networks through artificial neural network-based surrogate model. *Applied Sciences*, 10(18):6476.

Zeng, G., Li, D., Guo, S., Gao, L., Gao, Z., Stanley, H. E., and Havlin, S. (2019). Switch between critical percolation modes in city traffic dynamics. *Proceedings of the National Academy of Sciences*, 116(1):23–28.

Zhang, M., Yang, X., Zhang, J., and Li, G. (2022a). Post-earthquake resilience optimization of a rural "road-bridge" transportation network system. *Reliability Engineering & System Safety*, 225:108570.

Zhang, N. and Alipour, A. (2020). Two-stage model for optimized mitigation and recovery of bridge network with final goal of resilience. *Transportation Research Record*, 2674(10):114–123.

Zhang, N. and Alipour, A. (2023). A stochastic programming approach to enhance the resilience of infrastructure under weather-related risk. *Computer-Aided Civil and Infrastructure Engineering*, 38(4):411–432.

Zhang, W., Dong, H., Wen, J., and Han, Q. (2023a). A resilience-based decision framework for post-earthquake restoration of bridge networks under uncertainty. *Structure and Infrastructure Engineering*, pages 1–16.

Zhang, W., Wang, N., and Nicholson, C. (2017). Resilience-based post-disaster recovery strategies for road-bridge networks. *Structure and Infrastructure Engineering*, 13(11):1404–1413.

Zhang, X., Mahadevan, S., and Goebel, K. (2019). Network reconfiguration for increasing transportation system resilience under extreme events. *Risk Analysis*, 39(9):2054–2075.

Zhang, X., Miller-Hooks, E., and Denny, K. (2015). Assessing the role of network topology in transportation network resilience. *Journal of Transport Geography*, 46:35–45.

Zhang, Y., Zheng, S., and Chen, Y. (2023b). Identification of key nodes in comprehensive transportation network: A case study in Beijing-Tianjin-Hebei urban agglomeration, China. *Transportation Research Record*, page 03611981231192994.

Zhang, Z., Ji, T., and Wei, H.-H. (2022b). Dynamic emergency inspection routing and restoration scheduling to enhance the post-earthquake resilience of a highway–bridge network. *Reliability Engineering & System Safety*, 220:108282.

Zhang, Z., Ji, T., and Wei, H.-H. (2023c). Assessment of post-earthquake resilience of highway–bridge networks by considering downtime due to interaction of parallel restoration actions. *Structure and Infrastructure Engineering*, 19(5):589–605.

Zhao, J., Ma, W., Liu, Y., and Yang, X. (2014). Integrated design and operation of urban arterials with reversible lanes. *Transportmetrica B: Transport Dynamics*, 2(2):130–150.

Zhao, T. and Zhang, Y. (2020). Transportation infrastructure restoration optimization considering mobility and accessibility in resilience measures. *Transportation Research Part C: Emerging Technologies*, 117:102700.

Zhong, R., Chen, C., Huang, Y., Sumalee, A., Lam, W., and Xu, D. (2018). Robust perimeter control for two urban regions with macroscopic fundamental diagrams: A control-lyapunov function approach. *Transportation Research Part B: Methodological*, 117:687–707.

Zhou, T., Fields, E., and Osorio, C. (2023). A data-driven discrete simulation-based optimization algorithm for car-sharing service design. *Transportation Research Part B: Methodological*, 178:102818.

Zhou, Y., Lau, B. P. L., Yuen, C., Tunçer, B., and Wilhelm, E. (2018). Understanding urban human mobility through crowdsensed data. *IEEE Communications Magazine*, 56(11):52–59.

Zhou, Y., Wang, J., and Yang, H. (2019). Resilience of transportation systems: Concepts and comprehensive review. *IEEE Transactions on Intelligent Transportation Systems*, 20(12):4262–4276.

Zou, Q. and Chen, S. (2021). Resilience-based recovery scheduling of transportation network in mixed traffic environment: A deep-ensemble-assisted active learning approach. *Reliability Engineering & System Safety*, 215:107800.