UNIVERSITY OF TECHNOLOGY SYDNEY

DOCTORAL THESIS

Multi-modal Public Transport Modelling under Traffic Disruptions

Author:
Dong Zhao

Supervisor:
Dr. Adriana-Simona Mihaita
Dr. Yuming Ou
Dr. Hanna Grzybowska

in the

School of Computer Science
Faculty of Engineering and Information Technology

August 1, 2024

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Contents

1.1 Background and motivation

Public transport, which earns a reputation for its large capacity and environment-friendly with much greater success than private transport, has been growingly embraced in cities. It works as both a supporter of private transport, which facilitates personal mobility and leverages the infrastructures by carrying many more passengers in one run, and a competitor, which seizes the road capacity and land usage from private transports. To raise the image of public transport, apart from promoting the idea of its economic, social and environmental benefits, an improvement in service quality becomes more significant. The expectation on public transport travelling is a call for a multi-modal public transport system, where different transport modes covering bus, train, metro, air, and ferry have been introduced in cities to distribute various trips at the same time, while passengers are free to choose any suitable or preferred mode and route for travelling. Facilitating multi-modal public transport networks maximises mobility for various travel requirements, and the linked transportation solution along one direction largely expands the connectivity between locations. These benefits allow people to think about travelling alternatives apart from private vehicles.

However, an unexpected service disruption on a multi-modal public transport network would significantly affect the movement just because of the large capacity of the vehicles and the interconnection between networks of different transit modes. Due to the bond between networks, the propagation of the impact would deteriorate not only the current network but also the networks of other transit modes. Such disruption propagation works as a cascading failure where the breakdown of a node in an interconnected system can cause the downfall of other nodes through links, and continuously the initial failure of a node or link becomes a widely-impact disruption. This is the reason why network disruption needs to be explored. Presently, such network disruption propagation problem belongs to a widely-impact disruption problem because the impact of such disruption ends up lasting for a long period of time within a large range of areas. The network cascading failure phenomenon is a problem viewed from a dynamic level. The initial disruption itself can be a large disruption problem that impacts a large area for a long time, as well.

The integrated multi-modal public transport networks connect each entity by links, and the study regarding these connections in a complex network is known as the complex network theory. Under the field of complex network theory, various transport networks have been assessed and analysed to capture their configuration-oriented features by considering the degree of nodes, which means the number of connections a node has with other nodes (see Yi et al., 2021). Much evidence shows that the degree distribution approximately follows a power law, which means that the change of the degree is proportional to nodes, known as scale-free (see PastorSatorras and Diaz, Guilera, 2003; Clauset, Shalizi, and Newman, 2009). The power law of the degree distribution followed means that a large number of nodes connect only a few links while only a few nodes connect with a large number of links, and the subset nodes tend to have a small degree. Therefore, the scale-free network comes with a hierarchy of failure tolerance, reflected by the phenomenon that the scale-free network has a higher failure tolerance if the failure occurs at low-degree nodes because the impact of the failure can be limited to a small scale due to the limited number of connections, while the rest of the nodes are still functional enough to maintain the level of services roughly. On the contrary, the failure of those high-degree nodes commonly results in a traffic disruption against the travelling in a complex multi-modal-public transport network as more other nodes will be influenced via the connections (see the article conducted by Ortuzar S. and Willumsen, 2011; Shen, Ren, and Ran, 2019; Shang et al., 2020. The property in the scale-free transport networks is one of the reasons why traffic disruptions are more likely to be found in complex transport networks, and it is significant to provide targeting solutions for releasing the impacts.

Transport disruptions fall into two major categories according to the severity of their impact: small and large-scale disruptions. Most of the small-scale disruptions are recurrent disruptions, while large-scale disruptions can be either recurrent or, more oftentimes, current traffic disruptions. Small disruptions refer to those situations when road capacity and traffic flow are reduced for a short period of time within a certain range of areas. These traffic disruptions are relatively easy to solve, and the impacts on the individual, economy and society are limited. Common reasons that lead to small-scale disruptions include single-car accidents, traffic accidents due to the rear-end, side-impact or sideswipe collisions, failure of single public transport and short-term or partly road constructions.

On the other hand, large-scale disruptions themselves often catch more attention due to the considerable consequences. The large-scale disruptions associated with traffic interruptions often last for a long time or repeat much time. The consequences of this type of impact on traffic cover a larger area with a higher cost to the socioeconomy or even with the psychological influence on individuals. Typical examples of large disruptions can be found when natural hazards occur, such as bushfires, earthquakes, floods, tsunamis and extreme weather, or artificial disasters occur, such as significant accidents that lead to major individual injuries and major property damage, infrastructure collapses, terrorist attacks, large-scale transport services failure or significant traffic congestions.

Apart from major disasters, a minor traffic issue on a busy road can lead to a slowdown in traffic for a large area and a long time. If the vehicles involved in a small incident are moved, the disruption to the flow of traffic can persist for a long time afterwards and spread to a large area (see Saberi et al., 2020). Small road maintenance works can also lead to a major and repeated disruption. Even if it is just a single-lane closure, it can significantly slow down the overall flow of traffic, especially during peak hours (see Zhao et al., 2022). This is due to the connection between transport networks. When transport networks operate near their capacity during peak hours, even a small disruption can quickly saturate the system, leading to congestion and delays. The multi-modal public transport network offers travellers higher freedom in choosing where, when and how to travel; however, a disruption in one mode can have a ripple effect across others. This leads to delays in one part of the network and can propagate to other areas due to interconnected routes and shared resources.

The large-scale traffic disruption that resulted from a minor incident still impacts our daily travel. Although a growing awareness regarding this issue has triggered a number of research contributions, the same obstacles, such as the natural process of a disruption impacts the transport networks, the mechanism of disruption propagating into different networks and the way to model the impact of disruption for management purposes, are still existing and requiring to be addressed.

1.2 Research problems

Different causes lead to different types of traffic disruption and relevant impacts on network performances, which therefore require different modelling approaches and treatments. The consequences of disruption depend not only on the type and scale of the disruption itself but also on the circumstance and environment it involves. This is because of the heavy interaction between the internal entities and the structure of the complex networks. As mentioned in previous literature, some studies presented the disruptions based on a single network, while others modelled them in a multi-modal network but only considered public transport modes without considering impacts from the private transports; some studies showed the disruption in a small close network, while others suggested displaying them in an open network. The different modelling background environments summarised a different model of disruption. This phenomenon also raises a further question: can we unify the modelling methods not only against different types or scales of disruptions but also against different networks? Otherwise, can we summarise the regularity of the impact of network structure on the consequences of disruptions that result from road disruption?

The complexity of the disruption problem can also be displayed in their complicated dynamic features. The change of traffic can be split into three major phases at a disruption, where the first phase includes the process of traffic flow dropping because of the base disruption, the second phase covers the process that the traffic maintains low while the interventions are required for the journey recovery, and the third phase describes the situation when disruption is controlled, and traffic is returning back to normal. In practice, the degradation or closure of the roads due to disruptions may also result in difficulty of resupplying to the damaged areas, which broadly impacts emergency responses and incident recovery. Such post-disruptions against recovery indirectly exacerbate the impact of the base disruption on the networks, which has been commonly found in large-scale disruptions against different transport modes' networks. However this post-disruption impact can hardly be found in existing modelling methods in the literature, majorly due to the lack of data. The mechanism of the disruption impact locally and globally, especially when multi-modal public transit networks are considered, is still requiring more estimation. With the unbalanced demand and supply and the inadequate synergy of different transit mode networks, network recovery and passenger journey recovery have become another major challenge.

In addition to the complexity, large-scale disruptions are commonly considered as low probability events. This lack of data makes such sporadic disruptions difficult to be qualified or regularised into a model. Such a challenge forced most research studies to establish their disruption models according to the natural rule, but how accurate are the models compared to reality? Are those models with the assumptions still workable against today's transport networks? The challenging questions have sparked many endeavours, and this research study is also one of these endeavours that attempts to address the questions from a new angle. Since the availability of incident log and traffic state data (including travel flow and speed), more empirical evidence can be found based on data-driven approaches in order to support the work of modelling disruptions.

1.3 Research objectives

To understand the disruption impact as well as the propagation on current urban multi-modal public transport networks, the focus on a single mode of the transport network is no longer sufficient. Modelling the multi-modal public transport networks and the movement relying on different services is highly required before the in-depth study of disruption, as the configuration of the network plays an important role in network performance under transport disruptions. The public transport modelling is often separately treated by modes as an open network with strong failure tolerance. In fact, if the modelled network is extended by integrating all available transport networks in an area, the integrated network could be seen as a close network whose failure tolerance may be less satisfied. Therefore, the modelling of multi-modal public transport that accurately reflects the real network infrastructures is required. Disruption modelling applied to the model of the transport networks is required to accurately reflect the impact on traffic states such as traffic delay, mean travel time, flow, density, etc. Such impact analysis relies on both empirical and statistical analyses that are obtained from conventional impact analysis algorithms and incident or disruption log data. By combining both empirical and statistical results, the mechanism of disruption impact can be comprehensively abstracted to establish a disruption model for assessing or proposing the disruptions and disruption impacts' control strategies.

The major objectives of this research study include:

- Provide accurate multi-modal public transport demand estimation that reflects travel patterns in urban areas;
- Model accurately the temporal and spatial impact of the different types and scales of traffic disruptions on networks and provides the simulation model which is capable of abstracting the relationship between the network and the travel behaviour against disruptions;
- Estimate the disruption impact and the impact propagation in a multi-modal public transport environment which are reflected by the change in traffic states and travel efficiency;
- Deliver the impact functions and optimise the disruption model by investigating real disruption, traffic states and public transport patronage data.

1.4 Thesis organisation

After the research introduction, a comprehensive literature review is presented in Chapter 2, where previous research studies regarding OD estimation for public transport and traffic incident modelling methods are summarised. The following research is divided into three subsequent chapters. The Chapter 3 is related to estimating the large-scale OD matrices for multi-modal PT in urban areas. The proposed method relies on a novel model calibration method using Entropy-weighting, which considers both the traffic characteristics (travel time, travel distance, and fare cost) as well as the graph topological features (connections, closeness and straightness). In Chapter 4, we introduce a novel method for dynamically modelling PT patronage patterns by incorporating the Fourier transformation to effectively reduce the influence of noise and enhance the accuracy and reliability of the model. This model combines analytical techniques with data-driven methods to help identify the impact of incidents on PT passengers. The aim of Chapter 5 is to capture the spatial-temporal disruption impact via a multi-modal transport simulation model. In this chapter, we focus on a mode shift impact modelling to evaluate the best mitigation strategies and employ different disruption impact indicators such as delay time, flow, density and travel time for identifying the impacts. Finally, in Chapter 6, we provide concluding remarks based on our research findings and outline directions for future research endeavours.

Chapter 2

Literature review

This chapter is based on an edited edition of the following article: Zhao, D., Mihaita, A.S., Ou, Y., Shafiei, S., Grzybowska, H., Qin, K, Tan, G., Li, M., Multi-modal transport modelling under various disruptions: a comprehensive literature review, Under review at the Journal of Accident Analysis.

Contents

2.1 Public transport Origin-destination matrix estimation

Public transport (PT), which earns a good reputation for its large capacity and environment-friendly configuration in comparison to the road transit modes, has been increasingly embraced in cities supporting transport modes to carry many passengers every day, as stated in reports provided by the Department of Infrastructure and Regional Development, 2016 and OECD, 2018. Moreover, different transport modes, including bus, train, metro, air, and ferry, have been put forward across cities to distribute trips effectively. The advantages of using multi-modal public transport networks (PTNs) maximise passengers' mobility for various travel requirements between any location.

However, an unexpected service disruption would significantly affect mobility due to the system complexity and the interconnection between different transit mode networks. A disruption may propagate in a cascading manner where the breakdown of a node in an interconnected network can cause the downfall of other directly linked nodes. In transport networks, a failure of a node or a link may cause cascading disruption. However, in the literature, the way the impact of a disruption is captured at a local and/or global level is scant. Most studies model and analyse the problem by looking at individual transport modes in isolation from the rest of the transport system instead of applying a comprehensive and systematic approach. With the unbalanced demand and supply and the inadequate synergy of different transit modes, both network recovery and passenger journey recovery are a major challenge.

PT modes are often modelled separately as an open network with strong failure tolerance. However, to understand both the impact of a disruption and its propagation in current urban multi-modal PTN, the focus on a single mode of transport is no longer sufficient. In fact, if the modelled network is extended by integrating all available transport networks in an area, the integrated network could be seen as a close network with lower failure tolerance.

PTN modelling is a critical component of PT management and planning. Quality modelling can accurately reflect the current traffic states and existing problems, and help identify possible improvements to enhance overall system performance. These benefits have attracted significant attention from the transport community bringing about a transport modelling approach including trip generation, trip distribution, modal split and trip assignment. This approach first appeared in the 1950s and is also known as the four-step model (see Mitchell and Rapkin, 1954). The model ventures to catch up with the speed of change in mobility demand and supply, and considers the aggregation of trips and travel choices to simulate the travel behaviour that matches the reality in the network. With the use of specific methods and algorithms for each step, researchers not only can estimate the dynamic traffic demand of the network at anytime but also predict the future traffic circumstances under various scenarios. However, the efforts fall short of all the achievements as most transport and traffic problems are still striking our daily routines through congestion and ongoing incidents.

Greater confidence in PT modelling relies primarily on advanced demand estimation technologies. By knowing when and where a trip will take place, which transit mode will be chosen as well as which route will be picked by users, researchers will be able to understand the general travel pattern and distribution in the network. This process is known widely as the Origin and Destination (OD) estimation problem. In transport modelling, the OD matrix per transit mode is widely used to explicitly display the number of trips originating from and visiting specific zones or places. Consideration of multi-modal PT significantly raises the level of difficulty not only because it expands the number of

OD pairs that need to be estimated but also because it introduces a new challenge towards the mode choice as well as the interaction between networks of different modes. These challenges request more data from the real world. They also require more attention on the estimation side of trip patterns that reasonably considers the interplay among trips that use different modes. However, the lack of sufficient data and the limitation of effective methods when forecasting and estimating this multi-model trip matrix have been a repetitive problem that blocked the way towards conducting quality demand modelling for multi-modal PT. It is challenging, especially in large cities with a large number of daily commuting trips, to obtain the OD matrix by directly merging the trips of a large number of individuals, either due to a high computational effort or to a privacy issue. To date, there is still a high need for a better modelling approach.

A great many publications have reported on the attempts to improve transport modelling methods. This chapter contributes to the current body of knowledge by providing a comprehensive overview of the most recent contributions in the field of PT and demand modelling. The review starts by presenting the existing methods for creating an integrated network representation that matches the real multi-modal PTN (Section 2.1.2). The methods of network modelling are categorised by the traditional graph theory (Section 2.1.2), the layer-based theory (Section 2.1.2) and the Supernetwork theory (Section 2.1.2). The second part of this literature review covers the demand modelling approaches used in the recent research studies (Section 2.1.3), where the methods are grouped into the gravity model (Section 2.1.3), historical data-based models (Section 2.1.3 and Section 2.1.3) and other data-driven modelling approaches (Section 2.1.3). Finally, based on the knowledge obtained from the previous literature, this section provides comments on limitations, gaps and future directions in multi-modal disruption management research (see Section 2.3).

2.1.1 PRISMA diagram for PT OD estimation

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) is the review method that has been applied when organising and analysing literature for this section Page et al., 2021. This method was selected as it is well designed and organised for reporting systematic reviews with objectives, and the clear checklist and flow diagram for review was handy in use. The major processes for reviewing literature are illustrated in Figure 2.1. The first step aims to identify the relevant literature from the databases and other resources, such as websites, organisation publications and citation searching, based on keywords covering multi-modal transport network disruption, disaster impact on the transport network, large scale impact on the transport network, accident impact on the transport network, transport network failure and public transport network cascading failure in the transport network. The databases accepted for this stage were:

- ScienceDirect
- Taylor and Francis Online
- American Society of Civil Engineers(ASCE) Library
- SpringerLink

The database was accessed through the University of Technology Sydney, and the publication data was limited between 2015 to 2022 (the same condition was applied when searching other resources).

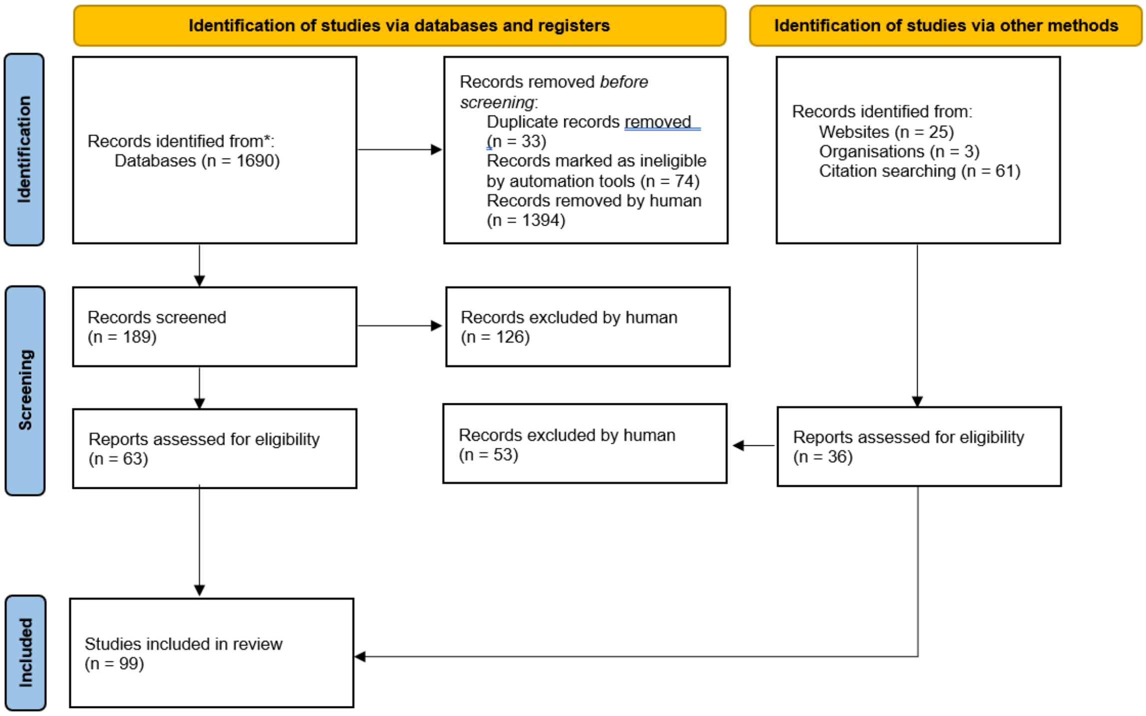


FIGURE 2.1: Flow diagram for systematic review based on the PRISMA approach.

From this process, 1779 pieces of literature were collected, where 1690 pieces of works were picked from databases and 89 works were found from other resources (25 articles were found on websites, three articles were published in organisations' official websites, and 61 articles were found based on citation searching). In this step, the selected literature was primarily screened by removing duplicates (33 works were removed), the ones marked as ineligible by automation tools (74 works were removed) and the ones not highly related to the subjects (1394 works were removed). The second step further removed 126 articles from databases, while 53 articles were removed from the database of citation searching because the articles were of a relevant low quality or out of the scope of this review article. The last step showed that the total number of literature that was ensured to be included in this review article was 99. For each article in the database, the initial review process started from the subsection of abstract, introduction and ended at the subsection of discussion or conclusion. Depending on the initial review, the highly related articles were identified as the systematic review article, and the secondary review covering the subsections of the methodology followed by the results was completed.

According to the literature, the focus on disruption analysis started in 2001, while most of the contributions were established between 2001 and 2013, followed by years between 2015 and 2017. Therefore, concerning this work's contribution, most attention has been made to the recent literature from 2015 to 2022, prioritised by the most recent works. Since 2019, more research studies have focused on the analysis of disruptions in multi-modal transport networks, where the agent-based transport simulation method was mostly used. Among all the literature, up to 48% of literature only focus on railway network disruptions, followed by a focus on road network disruptions which account for 22% and only 17% of research studies considered more than one transport modes network when analysing the traffic disruptions.

2.1.2 PT Network modelling

Graph theory-based modelling approach

The graph theory is a study of graphs, which aims to describe and help to understand the structures created by the relationships between the elements of a network, such as networks with multiple topological features. With the development of the graph theory, an increasing number of elements as well as the complex connection and interaction between elements are able to be illustrated using just nodes and edges. Such a study of graphing and analysing complex networks further delivers a complex network theory. The complex network theory first appeared as a speculation in the analysis of computer, biology, power grid, brain, climate and social networks before attracting the attention of transport networks experts.

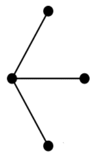
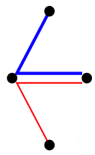
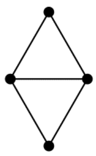
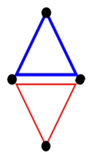
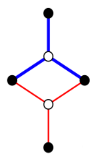
In the early days, the transit networks tended to be established in the form of either lattice (by increasing its randomness) or a random graph (see Renyi, 1959). With the development of graph theory, Watts and Strogatz, 1998 introduced for the first time the small-world structured network that had randomness between regular and random networks. After the analysis of three different real-world network architectures, this research concluded that all studied networks have a high vertices clustering tendency and low characteristic path lengths. The high vertices clustering tendency means that nodes in the network tend to gather together, and the low characteristic path lengths mean that the length of path connecting any pair of nodes in the network tends to be small. Following this, the clustering coefficient and the average path length have become the main indicators of the small-world network.

Followed by the small-world network concept, a scale-free network had been proposed and explained by Barabási and Albert, 1999. The scale-free network was named after the authors as the Barabási–Albert (BA) model. This research provided some empirical evidence indicating that the interaction between vertices decreases following the power law in the independent and large networks instead of the Poisson distribution, which had been commonly used in previous research studies. This feature was derived from a finding that most of the nodes are only connected to a few nodes (i.e., low connectivity), yet very few of the nodes are connected to an extremely large number of other nodes (i.e., high connectivity). The scale-free phenomena have two derived features: growing character (network tends to be extended by adding more nodes) and preferential attachment (new nodes tend to be added and connected to old nodes). These two breakthrough models, the small-world and scale-free network model, triggered the development of complex network theory-based modelling and analysing on transport networks, and further enabled the study of the impacts of network structure on both network dynamics and elements interactions.

The study of graphical representation directs the research studies to topological analysis and network analysis in most of the relevant publications; however the details regarding the following network measurement and analysis are out of the scope of this section, interested readers can find more references in literature by Lin and Ban, 2013 and Shanmukhappa et al., 2019. The Section 2.1.2 of this section aims to conclude the graph theory-based network modelling approaches that appear in the most recent publications, and summary the reason why such modelling approach is selected in specific research study.

In line with the study of complex networks, much evidence has shown that the configuration of urban transit networks presents itself as a complex network, where the local connection and non-local interaction among components in a system strongly interact with each other (see Regt et al., 2018). Such statements thus have been followed in order to build an abstraction of a real transport network, including bus, metro, urban train, light rail and multi-modal public transport networks in Austria, Australia, Brazil, Canada, China, England, Greece, Switzerland, Vietnam etc. The mentioned findings facilitated the development of network modelling under graph theory.

TABLE 2.1: Summary of graph theory-based network representation considered in literature

Network representations	Graph example	References
Space L		Space L or L-space graph is a topological graph consisting of nodes that are connected by a single edge if there is at least one existing path between them. In public transport networks, the PT stations or stops are represented by nodes, while the existing paths between two adjacent stops are represented by edges. The edges only connect node pair that belong to the same network.
Space L'		Space L' is derived from Space L, where multiple edges are allowed to represent different routes instead of connecting two nodes by a single edge like in Space L. In a Space L' graph, different edges represent different links used in different networks or the links with different directions in the same network, which help researchers to capture the inter-network properties.
Space P		Space P is the graph in which the nodes are connected by an edge if they can be reached from other nodes by services in the network; all the nodes of one line are connected to each other to show accessibility, even though there is no direct path. This means that the Space P graph also includes the transfer information and such graph is able to illustrate a multi-modal transport network.
Space P'		Similar to the definition of Space L', Space P' is the graph that includes edges from different networks or edges that are classified by structure characteristics such as directions.
Space B		Space B graph, also known as the bipartite graph, is a type of graph that uses nodes to represent both the paths and the PT stations, while the edges are used to connect the linked paths and stations. In Space B, the nodes are split into path or route nodes and station or stop nodes, and each station node is connected to multiple path nodes that it serves. In this type of graph, the black nodes are PT stations, while white nodes represent paths between each pair of stations, and the edges represent the connecting relationship between routes and stations.
Space C		In a Space C graph, the edges are used to connect two routes if they both link to a PT station. The nodes in Space C represents links, thus the nodes is also known as route-nodes in such space-based graphs. This type of representation only abstracts the relationship of routes from the network. Therefore, Space C is the most simplified graph representation.

According to the literature concerning graph theory and complex network theory, the transport

TABLE 2.2: Overview of the recent research studies considering graph theory-based network modelling

References	Layer objective	Type of network analysis	Weight objective	Space	Case study
Dimitrov and Ceder, 2016	Bus, train, ferry	Static	Unit demand-weighted edges	Space L	Auckland,NewZealand
Shanmukhappa, Ho, and Kong, 2018	Bus	Static	Bus flow-weighted edges; demand-weighted nodes	Space L	Hong Kong, China; London, UK; Bengaluru, India
Regt et al., 2018	Coach and train	Dynamic	Mean shortest path distance-weighted edges	Space L	Bristol, Manchester, West Midlands and London, UK
Luo et al., 2019	Tram	Static	In-vehicle travel times and waiting times-weighted edges	Space L	Melbourne, Australia; Vienna, Austria; Milan, Italy; Toronto, Canada; Budapest, Hungary; Zurich, Switzerland; Amsterdam and The Hague, Netherlands
Abdelaty2020	Bus	Static	Bus frequency-weighted edges	Space L	Cities in Canada
Du et al., 2020	Metro	Dynamic	Passenger volume-weighted edges; centrality-weighted nodes	Space L	Guangdong, China
Huynh and Barthelemy, 2021	Bus	Static	Average travel time-weighted edges	Space L	Ho Chi Minh, Vietnam
Zhang and Ng, 2021	Train	Static	Shortest path travel time-weighted nodes	Space L	Hong Kong, China
Kopsidas and Kepaptsoglou, 2022	Metro	Static	Travel time-weighted edges	Space L and Space P	Athens, Greece
Yang et al., 2014	Bus	Static	Minimum geographical distance-weighted edges	Space L’ and Space P’	Beijing, Shanghai and Hangzhou, China
De Bona et al., 2016	Bus	Static	Unweighted	Space P	Curitiba, Brazil; Beijing, Shanghai and Guangzhou, China; GOP, Warszawa, and Ł´od´, Poland
Zhang et al., 2006	Bus	Static	Unweighted	Space B	Beijing and Yangzhou, China
Sun, Lu, and Lee, 2015	Bus	Static	Overlapping bus stop-weighted edges	Space C	Singapore

network representations consist of various elements, such as nodes, edges or interactions between nodes and edges. An emphasis on different elements or interactions across elements generates distinctive directions toward network structural and topological characteristics analysis, which requires a targeted network abstraction when interpreting. To date, seven network representations have been found to best describe transport networks using the graph theory, namely, Space L, Space L’, Space P, Space P’, Space B, Space C and Space R, as summarised in Table 2.1, and have been used in the publications, see Table 2.2.

Layer-based modelling approach

Network modelling and representation have been largely improved thanks to continuous developments in digital, mathematical and geographical modelling. Concurrently, the role of network modelling has been maintained: to accurately represent the real-world transport networks. Due to the need of displaying a network's geographical structure, layers containing different structural elements of the network have been chosen to separate the network information and abstract the feature of each network element, which are normally symbolised by vertices, lines or polygons. Different layers contain distinct elements - this is the initial and essential concept of the layer-based modelling approach, it also an non-trivial extension of graph-theory network modelling approach. A variety of work has been done in relation to layer-based network modelling and analysis, which is summarised in Table 2.3.

TABLE 2.3: Overview of the recent research studies considering double or multiple layer-based network modelling

References	Layer objective	Type of network analysis	Weight objective	Space	Case study
Gallotti and Barthelemy, 2015	Rail, coach, air, ferry for inter cities; bus, metro and rail for inner cities	Static	Minimal travel-time weighted edges	Space L	UK, London, UK
Feng et al., 2017	Metro, metro service and passengers	Dynamic	Passenger flow-weighted edges	Space L '	Beijing, China
Yildirimoglu and Kim, 2018	Bus and passengers	Dynamic	Unweighted	Space L	Brisbane, Australia
Sui et al., 2019	Bus, passengers and transfer	Dynamic	Passenger flow-weighted edges	Space P	Chengdu and Qingdao, China
Tang et al., 2021a	Bus and metro	Static	Passenger flow-weighted edges	Space L and Space P	Shenzhen, China
Yang and An, 2021	Bus and metro	Dynamic	Passenger flow-weighted edges	Space L	Qingdao, China
Liu et al., 2022b	Bus and metro	Dynamic	Passenger flow-weighted edges	Space L	Xi'an, China
Pu, Li, and Ma, 2022	Bus and metro	Static	Unweighted	Space L	Lanzhou, China

The idea of using a layer to store geographical information has been soon extended to traffic information by appending more layers that describe transport users, transport services, traffic services or management. Following this, a layer-based framework model that separates activities, transport services, and traffic services have been firstly introduced by Schoemaker, Koolstra, and Bovy, 1999. Binsbergen and Visser, 2001, Nes, 2002, Kivelä et al., 2014 and Boccaletti et al., 2014 further investigated the transfer modelling, dynamic modelling, in-depth network structure analysis regarding the network robustness and the resilience towards private cars, taxis, freight transports, metros, trains, buses, trams and airlines networks.

More recent research studies were largely inspired by the concept of multiplex network modelling.

For example, Aleta, Meloni, and Moreno, 2017 compared two network modelling schemes: (1) layering by the PT lines, (2) layering by the transport modes, including walking, bus, tram and metro. However, this graph representation still ignored the interplay between road transport and the PT.

To understand and model the interaction and relationship between road transport and PT, Yildirimoglu and Kim, 2018 applied the Voronoi polygons to create small scale zones to cluster and unify the OD of the road transport and the PT trips (Von Landesberger et al., 2016). The road transport data was obtained from a Bluetooth data collector rather than a road path or link-based detector. The lack of OD locations forced the OD matrix to be built by self-defined small cell-like zones instead of residential or business areas following reality. The unification of the OD matrix for different means of transport can hardly be achieved.

Another improvement in the modelling resulted from the utilisation of the real-world data, e.g.: the vehicle GPS data allowing for tracking bus trajectories, and the smart-card data recording time, location and number of passengers boarding and alighting a bus. In this context, Sui et al., 2019 applied a layer-based modelling approach, but the road transport demand was ignored, and the bus network was represented by Space P (Figure 2.2).

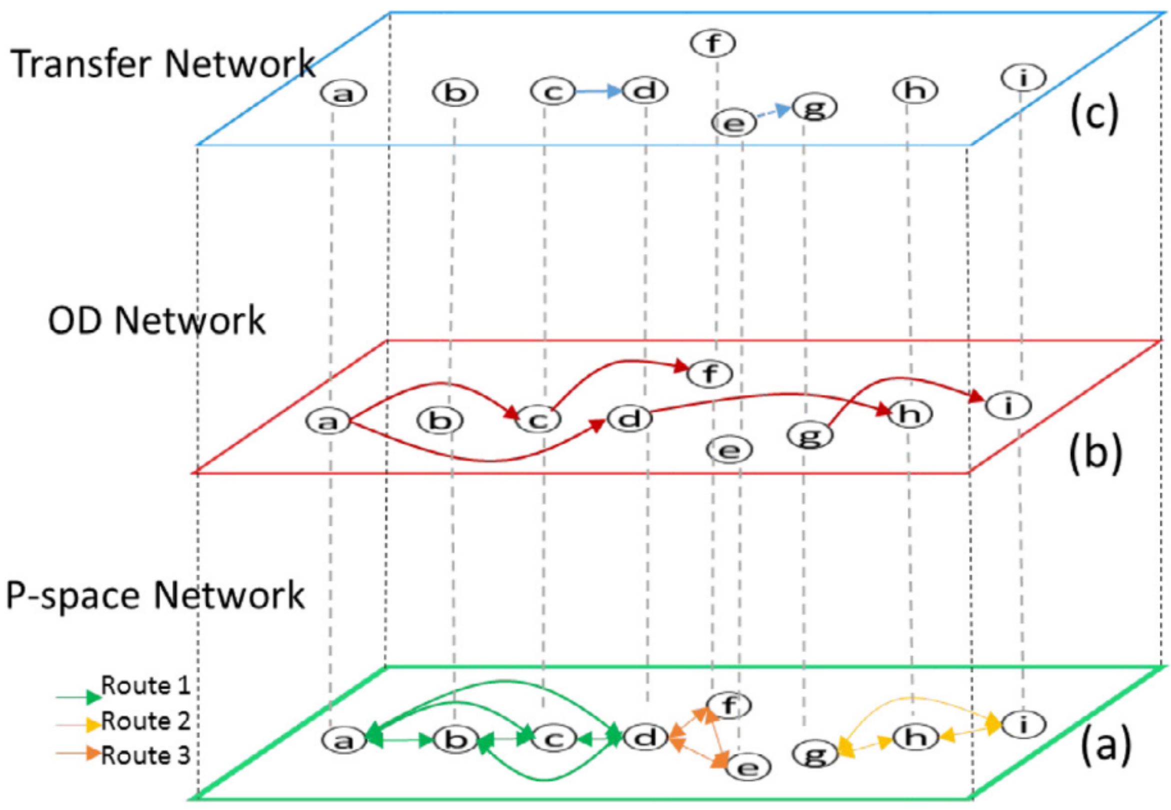


FIGURE 2.2: Description of the three layer model [Source: Sui et al., 2019]

Most recently, Huang et al., 2021 decomposed the bus system into three layers: (1) goal layer displaying a scenario when a passenger travels to their destination under several specific conditions (e.g.: good or bad vehicle condition, driver condition or bus schedule), (2) event layer describing the internal or external influence factors on movements (e.g.: road condition, bus operation reliability or weather), and (3) treatment layer defining control strategies that improve the level of bus services. This study put forward a comprehensive layer-based modelling that attempted to include as many elements in a bus system as possible.

Supernetwork approach

Although it is capable of replicating the travel behaviour based on the trip-based modelling theory, as represented by graph-based modelling method and layer-based modelling method, the extended interests toward the diversity activities that trigger travelling still call development in transport modelling. To cope with this, an emerging activity-based model comes with the potential of capturing the individual travel behaviour with a time-space dependence. The activity-based modelling derives from the activity or the need of travellers. The need of travellers can also be explained as the travel purpose; note that in activity-based modelling, it is the travel need that leads to the action of travelling. It is important to have this logical connection in mind as this is how trips are inferred in such modelling. Following the activity choice, the next challenge should be to model how people respond to the activity participation. This step corresponds to the choice of the activity location, the time that travellers should arrive at the activity location, as well as the mode choice for a movement of an individual. Since the whole process is derived from a single activity, which is the requirement for travelling, this modelling theory is also known as a heuristic demand modelling in some literature reviews.

The activity-based modelling is a response to the time-dependent trip (or trip chain) for an individual; the structure of modelling is similar to the traditional trip-based modelling: activity generation, trip destination choice, mode choice and trip assignment. The only difference with traditional trip-based modelling is that the Supernetwork considers a multilayer network framework instead of a single network. As shown in the Figure 2.3, this is a Supernetwork representation of parking and picking-up, the node of PVN represents private vehicle networks, and the node PTN represents public transport networks. The arrow lines indicate the activity between two nodes, in this example, the arrow line from PTN to PVN represents a picking-up process, while the arrow line from PVN to PTN means parking process.

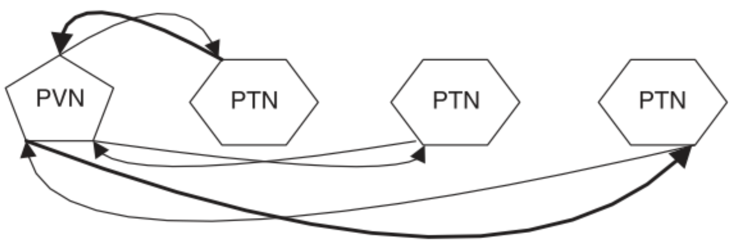


FIGURE 2.3: Example of parking or picking-up links by Supernetwork representation
[Source: figure taken from Liao, Arentze, and Timmermans, 2010]

The improvement in modelling for integrated transport networks has awakened the proposal of Supernetworks since Sheffi, 1985, where the full-scale modelling on multi-modal transport networks can display the travel patterns and behaviour within a complex system. Inspired by the Supernetwork concept, a model that integrates different transport mode networks into one single network has been proposed to uncover the interconnection among all modes, see the work done by Benjamins, Lindveld, and Nes, 2001. In this model, after generating potential routes for each population group who head to their favourite activities, a preference weight used in the Paired Combinatorial Logit model is applied for a route choice of each population group. The choice model was replaced by a heuristic model in the research of Arentze and Timmermans, 2004 to represent the process of decision-making.

The concept of day-to-day travel behaviour modelling following the Supernetwork representation has been proposed since Arentze and Timmermans, 2009 in order to capture the overall travel behaviour within a day. This was the initial attempt of using a microscopic modelling approach to capture the network dynamic feature. Liao, Arentze, and Timmermans, 2011 put forward to the travel behaviour modelling for a single person by using a heuristics rule related to the heading activity and the travel cost. Such heuristic models have been further strengthened into a time-space-dependent model by joining the time constraints so that the model is able to define the scheduling behaviour for each trip; details can be found in the research of Liao, Arentze, and Timmermans, 2013 and Li, Lam, and Wong, 2014.

Due to the consideration of time constraints, the computing time for travel behaviour and trip assignment at each timestamp have challenged the modelling effectiveness for the rest of years. Such an issue forced the insight on trip modelling to dispel the enumeration approach at a static level. One initiative on reducing the computing time was provided by Ouyang et al., 2011. He suggested utilising the multinomial logit function when calculating the shortest path at each interval of 10 minutes to speed up the convergence. Lam and Yin, 2001 presented an activity choice model based on the multinomial logistic model where the route was selected as an ideal dynamic user equilibrium model with the Method of Successive Averages (MSA). Motivated by this, a refined Supernetwork-based trip model was proposed by Fu and Lam, 2013 to capture the travel behaviour among cars, buses and subways; the process of mode choice was decided by a bi-criteria related to the mean utility and the budget utility that determines the travel time uncertainty at each iteration. The trip assignment model was build based on a dynamic user equilibrium using the MSA algorithm. The activity-travel assignment problem was attempted to be addressed as a discrete time-dependent event by Liu et al., 2015. However, the previous research results were only showed the predictability of activity-travel and the case study were all conducted based on a small-scale dummy network considering limited and simple entities such as trips to home, work, shopping and eating.

Summary for PT network modelling

The graphical representation of the transport networks has a long and rich history because how accurately the networks are displayed is heavily associated with the network analysis. As a first step of transport modelling, the transport network representations are also related to further traffic movement reproduction. Current contributions on the network representations when using the graph theory are able to flawlessly abstract the structure of a single transport mode network and comprehensively analyse the network structure by multiple indexes. However, the interconnection between different transport mode networks is still underestimated by a simple graph theory representation. This requirement resulted in an introduction of the layer-based modelling approach. The Layer-based transport network representation extended the modelling objects to factors that are out of the network structure and its properties. Such modelling method dimensional appends several graph-based layers into a comprehensive transport system modelling that includes physical transport infrastructure, stakeholders and services. Between the layers, there is also the inter-layer connection or relationship that can be magnified and modelled to better represent reality. Meanwhile, the activity-based modelling approach appeared in order to highlight the individual travel behaviour. This modelling approach turned the attention to a microscopic level where the interaction between different networks can be reflected by the travel behaviour of an individual and the travel cost spent on transferring, parking or picking up. With the development of computing and information technology as well as the real data covering vehicle location and demand dynamic, more and more network modelling were generated based on simulating approaches.

2.1.3 PT demand modelling

Gravity model approaches

The purpose of demand estimation is to understand the current and future usage of transport modes so that the designed transport management and operation schemes can be leveraged well to fit the demand. Therefore, techniques in estimation modelling became a critical task that has been attracting attention since the last century. Among the research contributions, the gravity model is noticeable to be introduced to capture the relationship between the trip distribution and the microscopic zonal demand. The development upon this estimation model and algorithm have appeared and can be classified by given data, normally, trip end data and link count data.

Modelling with production and attraction data In the study of travel patterns, there are generally two basic configurations: one is the OD estimation; the other one is the Production to Attraction (PA) estimation, as described by Ortuzar S. and Willumsen, 2011. The OD matrix contains several OD pairs, which equals the total number of trips from a specific location to another. Such representation can be aggregated with respect to the type of users, travel purpose or transit modes. The PA matrix refers to the attracted trips from a production place. Taking a journey to work and back during a day as an example, two trips are produced: one trip happens in the morning, which is a trip originating from a home point, and the other one occurs in the afternoon, which is a trip generating from the workplace point; there are also two trips that are attracted by both home point and workplace point on that day. While in terms of the OD matrix, there is only one trip from home to the workplace and the other trip from the workplace back to home.

Research in investigating the travel patterns and behaviour based on OD estimation dates back to the period when a gravity-based model was introduced to describe the relationship between towns and retail trade (see the work of Reilly, 1931). Such application of the gravitation law triggered much research to focus on the trip estimation between a pair of nodes, as introduced by Dieter, 1962 who built a model based on the gravity law and clearly indicated that the number of trip between two area was proportional to the total number of trips at origin and destination, which was the total number of production and attraction, and revering proportion to the distance between groups.

Considering to solve the gravity model, Bouchard and Pyers, 1965 and Ben, Bouchard, and Sweet Jr, 1965 treated the model as a singly constrained model and only the zonal attraction parameter was considered to ensure that the modelled number of trips attracted by a zone matches the survey result. The main drawback of using such an abridged constraint was that the model can hardly ensure that the calculated total trips between an OD equal the sum of the trip departing and trip arriving. This limitation encouraged the practice of the doubly constrained gravity model, where both parameters for trip origins and trip destinations are considered (see Evans, 1973.

Modelling with link flow data Considering the difficulty when obtaining the travel production and attraction data for private vehicles, it is much realistic and available to track the link count for estimating the car demand. This estimation method was treated as the converse or iterative process in the trip distribution and assignment processes, while the relationship between estimated and observed flows could be fitted by a logistic function, which represented the compassion between estimated and observed results for calibration purposes. This concept that allies to the traditional gravity model have been shown in the research based on link count data instead of travel attraction and production data since Low, 1972. The model built by Low in 1972 started by collecting a large number of individual trips in a roadside interview survey as the initial link volume data. The estimated link volumes were calculated based on a format of the traditional gravity model subjecting to zonal population, employment and travel time instead of using the total number of trip production and attraction. The estimated link volumes named trip probability factors in this research were used as the travel demand in the further trip assignment process. The trip assignment ended up with the assigned link volumes, which were compared with the observed volumes for validation purposes. Thus, the relation between assigned volumes and observed link volumes could be expressed as a function generated by the regression program. Such an approach introduced a data-driven way to calibrate a postulated gravity model from the population, which initially put forward for estimating OD matrix from link count data. The model has been further developed with the consideration of car possession and social-economic level, as detailed in the book of Organisation for Economic Co-operation and Development, 1974.

More continually works focus on the solution of a mathematical equation which described the relationship between estimated traffic flow and variables that can be observed. Robillard, 1975 attempted to represent this relationship by using both non-linear least square estimation and the linear regression least square estimation. Högberg, 1976 has been inspired by Low's model (see Low, 1972) and provide a optimised non-linear regression model and addressed the regression model by the least-square algorithm.

Most of the previous research studies relied on artificial data and toy networks when validating the methods of OD estimation. The use of fake networks largely reduced the impact of network complexity on trip distribution and assignment. To fill the gap, Willumsen, 1981 systematically summarised the previous work regarding OD estimation from link count data and provided empirical evidence using real-world link flow data and a small range of real networks to verify the model feasibility. The estimation method used in this research was entropy maximising.

The most recently the attention has focused on the provision of the OD matrix when struggling to cope with very limited datasets. Research conducted by Thompson et al., 2019 investigated the input data that included a part of the population, the distance of links between population nodes (cities or towns as defined in the United States Census Bureau) and junctions, as well as the directional link flows. Without the total production and attraction data, this research further developed a cumulative gravity model to determine the path-based traffic flows between two population nodes. Such a model uses the relationship that the sum of the traffic on the path connecting a population node and its adjacent population nodes equals the traffic between this population node and its connected junctions.

Estimation of the deterrence function Along with the development of the gravity model, the influence factors that reversely yield the total number of trips between OD pairs attracts much attention. Rather than using a single variable such as travel time, the element was further replaced by a deterrence function that is broadly composed of a group of travel costs, such as travel time, distance or trip length, the average travel speed and monetary cost, as summarised in the research of Erlander, Nguyen, and Stewart, 1979 and Sen and Smith, 1995. Since then, the configuration and parameter calibration has been known as a key condition when estimating the OD matrix by many researchers.

A typical "guessing" approach was applied by Hyman, 1969 who proposed a robust method triggered by Bayesian estimation in 1969. The author was triggered by his experience that the initial deterrence function parameter should be between 1 and 2 and proposed the equation to express the relationship between parameters and total travel cost:

$\beta = \frac{3}{2C}$,

where β is the initial parameter of the deterrence function and C is the total travel cost. The initially calibrated deterrence function could be employed at the trip distribution process. The new OD matrix, as the output of this initial trip distribution, generated a new total travel cost distribution which was used in a new cycle of deterrence function calibration. Following this process, the best value of the deterrence function parameter could be found when the proportion of the total trip origin and total destination equals the proportion calculated based on the observed data, and the mean trip cost of estimated data equals the observed one.

Influenced by Hyman, 1969, Evans, 1971 adopted the Exponential function as the decreasing function and used the Maximum Likelihood estimation to calibrate the function. The main idea of this approach was to find the parameter that maximised the probability of a situation when the estimated mean trip cost equals the observed one. This research then suggested using the Chi-square test for explaining the goodness of fitting, but there was no numerical evidence provided in this research.

Unlike those methods started by guessing a parameter for the deterrence function or iteratively calibrating the functions, the method of Evans, 1971 directly calculated the parameter by solving a polynomial function. The approximating polynomial function was solved using Maclaurin Seris expansion about the deterrence function parameter, and such calculation process was also explained with a toy numerical example. The Maximum Likelihood approach was proven to be reasonable mathematically, but the workload required by the calculation limited its application on calibrating the deterrence function from a large amount of OD trip costs.

Hathaway, 1973 also practised the Maximum Likelihood estimation to find the most probable deterrence function parameters. He assumed that the likelihood function of the deterrence function parameter was approximately normal. Then the maximum likelihood could be found at the mean of parameters. This method intended to find a value that approximated the value of the deterrence function parameter, whereas the method suggested by Evans, 1971 straightly calculate the best value of the parameter for the deterrence function. Therefore, the accuracy of the parameter calculated by Evans seems to be relatively higher. Further comparison among Hyman's method, Evan's model as well as Hathaway's estimation can also be found in the work of Williams, 1976.

The idea of using only the total travel cost in finding the parameter of the deterrence function had been proven unreasonable since Erlander, Nguyen, and Stewart, 1979. The research study conducted by Erlander in 1979 embedded the trip distribution and assignment into one model, instead of using the total travel cost, the deterrence function chosen in this research was a component of two monotone functions that subjected to travel cost and entropy of the travel demand. The configuration for the travel cost-oriented monotone function was still an exponential function. In this case, the gravity model could generate the link flow while maintaining a minimum total travel cost and maximum travel demand entropy before the equilibrium flows were assigned to the network. The research had found that the unique parameter of the deterrence function could be found when using the combined monotone functions, while the use of only total travel time was proven insufficient to calculate a unique value for the parameter. Yet, only the mathematical formulation-based proof was provided in this research, and there was still a need for further numerical evidence.

The expression of the travel cost was changed by using a form of the gamma function in a research conducted by Daly, 1982. The deterrence function was calibrated using a linear regression algorithm to find the result when the gap between the estimated total trip number and the observed one was equal. This research also tested the impact of model size on the effectiveness of the deterrence function calibration by reducing the number of available destinations. According to the numerical evidence, this research stated that, compared to the result run with 20 zones, the result run by eight zones had more standard errors. Further research or more numerical comparisons are still needed before a certain statement on the impact of a number of zones included in the modelling on calibration effectiveness, as a comparison of modelling by 8 and 20 randomly chose zones is inadequate to ensure the relationship.

A form of the power function subjecting to travel distance in research later carried by Shen and Aydin, 2014, In Shen's research, the gravity model was established to estimate the OD matrix for highway freight transport. Similar to the calibration method applied by Daly, 1982, the least square regression method was used to find the best parameters that enabled the model best interpret reality.

While Abdel-Aal, 2014 decided to use the Exponential function when modelling the impact of travel distance in deterrence function to trips with different travel purposes. The difference between the research of Abdel-Aal, 2014 and previous research studies was the extension of including the trip purpose for each zonal travelling: the deterrence function parameters were split and defined based on different trip purposes. An Iterative Proportional Fitting (IPF) approach was employed in this research for model calibration. This approach is similar to the "guessing" approach introduced before. However, this study started to use 0.1 as the initial parameter of the deterrence function, and the best parameter was iteratively found when the estimated trip length distribution equalled the observed one.

The concentration on gravity modelling was shifted to the public transport by Tamblay et al., 2016 who inferred the OD matrix for trips by bus and metro from a stop-to-stop based smartcard data. This research study initially considered not only the zonal information embedded into the trip departure and arrival but also the public transport stop information. Drawn from Daly, 1982, the deterrence function was established by two gamma functions subjecting to zone-to-stop distance and stop-to-zone distance, separately. The Maximum Likelihood estimation directly calculated the best-fit parameter according to the observations. Such a method saved much more time than the iteration-based trip assignment method and, therefore, was largely approximated to reality.

A most recent research study (see Thompson et al., 2019) that focused on freight transport used the Power function to describe the impact of travel distance on trip distribution. Given the link count data, the parameter of the deterrence function was calibrated by the Least Squares Regression, similar to the method used by Daly, 1982 and Shen and Aydin, 2014. One more step that was applied in the research of Thompson et al., 2019 was the test for parameter fitting. Common Part of Commuters measure (CPC) and Coefficient of determination (R square) were used as a further step to test how well the observed data were replicated by modelling.

Before using a model, it is always essential to calibrate the model. The gravity model that describes the relationship between the number of directional trips and the total trip end information significantly rely on the natural travel cost. Derived from this feature, a calibration of the deterrence function is as vital as calibrating the parameters of the gravity model. Much previous research proposed different approaches based on their given data, primarily smartcard data or survey data. However, there is a gap in finding a way that used the minor real data to achieve the purpose of deterrence function calibration. Most studies failed to care for the configuration of the deterrence function as well, but directly using the most common form of function. This neglection primarily due to the limit of real data.

Maximum entropy approaches Another approach that has been instrumental in generating a range of models to solve the trip distribution problem, such as the gravity model and market model, is the entropy maximisation approach, as detailed by Murchland, 1966. The proposed method was derived from the theory that a set of distributed trips that best represented the reality should have the largest entropy. Therefore, the objective function was extended to a combined function of entropy maximisation and travel cost minimisation, and the best parameter of the gravity model could be iteratively determined by trip distribution entropy. Such the solving method have been proved fitability and versatility by Wilson, 1967; Wilson, 1969; Wilson, 1971.

Fang, Science, and 1995, 1995 proposed an optimised entropy-based trip distribution model by considering a convex quadratic cost to better describe the impact of congestion on travel cost. Real survey data (total trip production and destination data) was used to calibrate the entropy-based model by Wang, Yao, and Jing, 2006. Xie, Kockelman, and Waller, 2011 applied the entropy maximum approach when treated the the travel demand as a elastic function to OD location. Li et al., 2011 further considered the situation when travel cost were fuzzy variables and the total trip production and destination were random variables, and proposed an advanced chance constrained entropy-based model. Tang et al., 2018 demonstrated the power of big data by employing taxi GPS trajectories data when calibrating the entropy-based model and proved its superiority in trip estimation. More recent research study conducted by Huang et al., 2020 considered the impact of time on trip distribution and proposed a spatial-temporal related maximum entropy-based model.

Direct demand approaches The direct demand model for an OD estimation derived from the demand theory and market model concept which was introduced by Kraft, 1962. The demand theory was originally prompted in economic principle that demonstrated the connection of demand for a good to its price in the market. When using in the field of transportation, the travel demand and supply of the transport services depends on the travel cost that always displays as the demand curve, where low travel costs attract travel demand while reducing supply, and high travel cost situation losses travel demand but increase service supply. Direct demand theory follows the tendency in the demand curve and creates demand estimation modes that are related to both the social economy and the development of infrastructures. Several early example of research into this concept has been undertaken, as reviewed in Stopher and Meyburg, 1975 and Gaudry and Wills, 1977.

The major logic of this method, as described in Carey, Hendrickson, and Siddharthan, 1981, started from randomly creating an OD matrix as the initial OD matrix and assigning the trip in the network by static user equilibrium for obtaining the link count data. The iterative calculation for trip assignment stops when the estimated link count data equaled the real link count data, and in this way, the research was able to deliver a relatively capable OD estimation that matched the historical real data. Carey put forward an approximation of the objective function towards the disparity between estimated flow and real flow using a linear function. Such the used algorithm was also known as the Frank-Wolfe algorithm. In his research, the objective function is solved by least square estimation. Since the model is built in a toy network consisted of a single route, the method proposed is heavily required to be extended for more general use. Therefore, the author included not only the cordon count data but also ticket count data for buses and Census data that involves the individual zonal journey to work information Carey and Revelli, 1986.

Following this work, Rios, 2001 developed a model for estimating OD matrix from link count data by using the direct demand model. His work improved the objective function with regard to minimising the gap between real and estimated data among origin trip ends, destination trip ends, sample OD matrix, total OD trips and link count, instead of the origin and destination trip end only. The model was adjusted to a non-linear function that was formed by both maximum entropy and least square functions to reduce the iteration duration while increasing the adherence to historical data. The solutions of the objective function introduced in this literature are the Z-approximation heuristic which is similar to the Frank-Wolfe algorithm, and the Delta-approximation heuristic which additionally considers the gap between previous and later iteration results of the OD matrix derived from the trip assignments. The author also highlighted the impact of input data on modelling results in another research Ríos, Nozick, and Turnquist, 2002. This research finding suggested that the input data mixed with 40% origin and destination trip end data, 20% sample OD pair data and 20%link flow data produces the most accurate OD matrix for any freight network. Kepaptsoglou, Stathopoulos, and Karlaftis, 2017 applied the direct model when predicted the ridership of light rail transport based on historical demand data and the OD data from a roadside survey. This research also provided a literature review about the application of direct demand model in rail transit system that covered the publications from year 2004 to 2014.

Since the mechanism of the direct demand approach is to uncover the relationship between social-economic and transportation attributes-related variables and the travel demand, the machine learning algorithms could be the superior options, as proposed by Yan, Liu, and Zhao, 2020.

Other data-driven approaches Over the past decades, studies have provided important information and techniques on OD estimation. However, most of the previous research used directional trips derived from the human resources-based OD survey directly, such as the endeavours contributed by Ben, Bouchard, and Sweet Jr, 1965, which is very similar to the work of Organisation for Economic Co-operation and Development, 1974. Moreover, the mentioned researches focus mostly on statical OD estimation that has ignored the influence of time on the OD matrix. The study on static OD estimation is principally limited by the raw data itself as the technology for obtaining time-dependent data is a challenge. But with the development of the data collection systems and technologies, it is able to have big data that is time-dependent now.

Modelling with historical OD matrix or link count data

The study of Yang, 1995 involved a generalised least squares (LSQR) estimation in calibrating the estimated link flow data from the observed one. The use of LSQR largely enables the simulation to approach reality and essentially eliminate the impact of the nature of underestimating. For the second level of the modelling, two heuristic traffic assignment models were tested. The first model referred to the iterative estimation assignment based on user equilibrium theory; the second one was a direct sensitivity-based OD analysis drawn from the research of Tobin and Friesz, 1988. Triggered by finding the best algorithm for model calibration, Cascetta and Russo, 1997 firstly compared the methods of maximum likelihood, non-linear generalised least-square and Bayes estimation when using observed link count data for OD estimation. The author stated that the non-linear generalised least-square and Bayes estimation are equally effective when addressing the link flow-based gravity model. However, since the non-linear generalised least-square is insensitive to the model misspecification and more flexible when defining, the mentioned research recommended the linear generalised least-squares to be the first choice. The rest of the research examined the feasibility of using the non-linear generalised least-square method for OD estimation from link count data. The method was tested based on real data in the multi-modal transport networks that involved the mode of pedestrians, cars and buses separated by travel purposes: home-to-work, home-to-school and home-to-other-activities. The objective function was iteratively solved by the gradient project algorithm of Rosen with linear constraints.

Nielsen, 1998 turned the attention to better assign trips in the network while iteratively calculating the OD matrix to smooth out the data error. In his research, the route search followed the prevalent Method of Successive Averages (MSA) and Dijkstra's algorithm, and the traffic assignment process obeyed the Stochastical User Equilibrium (SUE) theory. Two different trip assignment models have been proposed in the research to demonstrate the influence of the different sets of observed link count data on the estimated result. The first method is Single Path Matrix Estimation (SPME), where only the optimised paths selected by the MSA is assigned with trips; this is analogous to the all-or-nothing method. On the other hand, the method called Multiple Path Matrix Estimation (MPME) considers all possible paths and weighs the path by probability to indicate the preference of path choice. The research proves the superiority of the second method though it takes more time to compute.

Having the time-dependent flow-based data, the vehicular OD estimation can be formulated stochastically at the trip assignment process. Time dimension largely increases the workload of calculation which triggers methods such as the Kalman Filter to effectively tackle the non-linear problem by using a first-order Taylor linearisation, which is further discussed in Ashok and Ben-Akiva, 2002. Moreover, research conducted by Bierlaire and Crittin, 2004 recognised the significance of time and the benefits of big data. Acknowledging the significance of time dependence in a demand prediction, this mentioned research attempts to capture the time-dependent OD matrix from the time-dependent historical OD matrix by using the least-square approach. Two widely adopted algorithms, namely Kalman Filters and LSQR, for reducing the gap between estimated and observed trips were compared in this research. Bierlaire indicated, based on the result from both theoretical analysis and the empirical experiment shows, that the LSQR tends to be better when solving the dynamic least-squares problems.

Research studies convoyed by Ashok and Ben-Akiva, 2002 and Zhou and Mahmassani, 2007 aim to optimise the estimated OD matrix by using Kalman filtering while iteratively assigning the trips based on the dynamic traffic assignment model. Cantelmo et al., 2015 splitted the OD estimation problem by applying a two-step optimisation method to produce an initial zonal production and attraction matrix that matches the real data through iteration before using the produced PA as input data in a dynamic traffic assignment model to generate the most likely OD matrix. The dynamic trip assignment model is built along with a superior path searching algorithm called Simultaneous Perturbation Stochastic Approximation (SPSA) which has a strong ability in calibrating large-scale traffic assignment model due to perturbing all the variables in the objective function at once, as reported in the research of Balakrishna, Ben-Akiva, and Koutsopoulos, 2003 and Lu et al., 2015.

More recent research carried by Xie, Kockelman, and Waller, 2011 attempted to obtain the OD matrix based on link flow data by using a new method combined entropy maximising and least squares estimation. The entropy maximisation was used at the trip distribution process, while the least squares was used to obtain estimated link flow from observed flows. The outcome of these models produced the flow rate for each link that could be used for link flow prediction directly rather than calculating the link flow by accumulating the trips through the trip distribution process. The demand function revealed the relationship between estimated link flows and observed link travel costs that is analogous to the non-constrained gravity model. Such a model on OD estimation converts the linear approximation problem to a computable combination of maximum entropy and the least squares problem, which enormously improve the effectiveness of calculation compared with the previous iterative trip distribution process.

The experience from recent studies has been further developed, and new insights into demand prediction from the view of machine learning were proposed by Krishnakumari et al., 2019. Given the historical OD matrix data, the author aggregated the data for zonal production and attraction before employing the gravity model to generate the logit route choice model. The deterrence function corresponds to travel time only due to the available datasets that this research has. The model describing the relationship between the OD matrix, the trip production, the trip attraction and link flows were solved by a constrained linear least-squares algorithm. The part of demand prediction in this research offers an algorithm that applies an artificial neural network (ANN) model to uncover the relationship between inputs (either individual speeds and flows or the average zonal speeds and flows) and outputs (trip production and trip attraction). The model training highly depends on the initial OD matrix, which is a challenge in future research where the OD matrix is an extravagant expectation. All above mentioned data-driven methods were achieved since the existence of proper datasets. Especially with the details of time, the modelling result has been increasingly developed towards reality. More attempts on demand prediction by machine learning can be found in the research studies of Yan, Liu, and Zhao, 2020; Ma et al., 2019; Ou, Mihaita, and Chen, 2020; Shafiei et al., 2021b.

Modelling with historical smartcard data

In recent years, research studies on OD estimation began to emerge along with the increasing use of smartcard data, where much accurate historical tap-on and tap-off information is available to capture the travel pattern. Such smartcard data brings transportation modelling up to another level. Since the possession of the OD matrix, apart from OD pair estimation, much research has been carried out in the model and the simulation calibration, as reviewed in Hussain, Bhaskar, and Chung, 2021.

Integrating the automatic vehicle location (AVL) data with automatic fare collection (AFC) data, research that aims to validate a citywide transit assignment model was conducted by Tavassoli, Mesbah, and Hickman, 2018. The proposed assignment model was effectively calibrated by adjusting the model parameters manually until the difference between the modelled result and real data reached its minimum. The construction of the assignment model is accomplished in EMME, and three embedded validation measures, namely R-square, root mean square error and GEH, are used to display the effectiveness of the simulation from the perspective of a number of boarding or alighting passengers at stops, on route travelling and by direction. Similar attempt was achieved by Ou, Mihaita, and Chen, 2020, who focused on the urban rail network and estimated the OD matrix from integrated production and attraction data by using traditional four-step demand modelling approach. This research study also initially introduced a method to uncover the in-station trajectory by applying the hybrid Markov model with a consideration of the OD matrix, vehicle occupancy as well as the real time train schedule.

Modelling with other types of data

The lack of initial data can hardly stop the researchers' passion for OD estimation. Even without access to ready-to-use OD data, apart from the mathematical gravity model, a traditional probe vehicle can be used to count the link flows (see the research generated by Yang et al., 2017). In the study of Yang, the least squares estimation was utilised to calibrate the estimated flow count data from count data. Note that in his link flow estimation, the weight of the link choice was produced based on the travel cost, which is calculated when the trip assignment-based iteration is in use for the OD pattern estimation.

Similarly to the method of a probe vehicle, other studies used the vehicular identification system to calculate for obtaining the Electronic Toll Collection Data, where the directional information on a toll road can be recorded and used for OD matrix generation (see Jiang and Warita, 2009 and Kim et al., 2014).

There is another type of big data approach derived from GPS that can be used for OD estimation: mobile phone data. However, before the step of an actual estimation towards the travel pattern, data processing is required to identify the production and attraction location (see Iqbal et al., 2014 and Alexander et al., 2015). The processed data is then classified into the home, work and other locations before generating a completed trip data based on PA. As the cell phone data only roughly provides the arrival time and the remaining time at a location, the time that an individual departing has to be determined from external survey data. In the research of Alexander et al., 2015, the daily number of trips is treated as a uniform distribution. The number of trips for a period of time is simply calculated by multiplying the possibility of travelling with the total population. To calculate the daily number of trips, the result was obtained by being divided by the number of the simulated day. The outcome of daily trips is reformed by hour and aggregated by weekday and weekend based on travel purpose, namely home-based work, home-based other and a non-home-based trip. For validation purposes, the observed trip chain was obtained from mobile phone records: the record of location pausing is the time a passenger spending on an activity and the time when the location of a passenger was changed was recorded as the departing time. Other applications of cell phone data can be found in a review by Chen et al., 2016 and more recently in a research study of Ge and Fukuda, 2016 and Gadziński, 2018.

Bluetooth data that captures the vehicle trajectories also becomes a source of the traffic data. As used in research of Yildirimoglu and Kim, 2018, the Bluetooth detectors were located at the inter subsections, arterials and state-controlled roads in order to record the passing vehicles. The vehicle could be identified by its unique ID so that the trajectory can be mapped from a time-ordered sequence of passing each Bluetooth detector. The problem with this data is that the real origins and destination location cannot be detected as the trajectories are only detector-based.

Summary for PT demand dmodelling

Extensive evidence considers the OD estimation from total production and attraction data based on the gravity model. Such a method has been largely developed with the improvement of mathematical, analytical and computational skills. The model has been applied from displaying the relationship between population and trade benefit to, nowadays, demonstrating the directional trips from aggregated production and attraction data. However, the ability of the gravity model in the transport field is still underestimated, as most research studies attempted to investigate the OD matrix for a single transport mode, mostly cars. There is still a need to consider the influence of other transport modes when mode splitting and trip assigning under a multi-modal public transport environment. The challenge against integrating the OD matrix of public transport with that of the vehicle is still out there.

Following the purpose of estimating effectively, more convergence criteria have been proposed, such as the maximum entropy and maximum likelihood. Such methods follow the nature of the state of a system where the most likely state of a trip assignment is the situation when the system entropy reaches maximum; and in terms of the maximum likelihood method, the most likely state of a trip assignment is the situation when it best matches the observed data. Both methods set the calculation process with clear stop criteria, and it is in this way that the calculation becomes effective and the result of the OD estimation becomes accurate. The rest of the progress on OD estimation emphasised capturing reality and interpreting it by mathematical equations. Either using mathematical models or data-driven approaches, they are both able to uncover the hidden pattern and trend of travelling in networks.

2.2 Traffic disruption impacts on PT

Transport is a critical part of people's lifestyles, especially as cities expand. To meet the growing and diverse mobility needs, multiple public transit (PT) modes such as trains, buses, metros (underground/subway), on-demand buses or cars, and micro-mobility solutions like bikes, e-bikes or e-scooters have been introduced to support a sustainable transport system. This new era raises concerns about the synergy between networks for both PT and private vehicles. While cooperation among different transit networks enhances mobility, it also introduces new vulnerabilities due to the increasing number of entities integrated into the original network. A multi-modal PT network allows for impact propagation, leading to large-scale disruptions due to its reliance on shared infrastructure and transfer points. Additionally, disruptions can alter passenger behaviour, increasing demand for alternative modes or routes and further weakening the network. PT schedules are often coordinated to optimise transfers between different modes, so disruptions in one mode can cause cascading delays across the entire network. Such negative influences produced in a multi-modal PT environment encourage an increasing number of research studies to focus on disruption causation, measurement and disruption modelling.

However, we did not find any studies that compile all related works or perform a bibliometric analysis of the literature in this field. Based on existing publications, this literature survey provides a research framework to outline developments in traffic disruption and impact modelling as well as the application of existing modelling and research methods. Therefore, in this literature review, we aim to provide a comprehensive examination of existing publications concerning PT disruption causation, measurement, disruption modelling methods and their applications. We also provide opportunities and future directions in the field of multi-modal PT disruption study that guide future works.

While conducting the literature review, we encountered several open questions:

- What is the development of research regarding traffic disruption on PT?
- What are the essence research perspectives involved in publications?
- What are the major causes and the impact of traffic disruptions on PT?
- What are the principal methods for measuring the impact of specific disruptions?
- What is the primary focus of disruption modelling methods?
- What gaps exist in current research studies?
- What opportunities and future directions exist for multi-modal PT disruption studies?

By addressing these challenges, contributions can be achieved by conducting this research:

- Applying a novel bibliometric approach to examine existing publications on the impact of traffic disruptions on public transport and identifying the evolution of disruption impact on the PT network;
- Classifying the primary causes and impacts of traffic disruptions on public transport networks and the modelling methods that quantify the impact;
- Specifying gaps and challenges and pinpointing opportunities for future research.

This review article is organised as follows. In subsection 2.2.1, we present the details of the article review process based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Based on the selected publication using the PRISMA method, we conducted a bibliometric analysis to reveal trends in related research, focusing on publication year, author, keywords, case studies, and literature sources. The causes of disruptions involved in multi-modal PT and the impact measuring methods are reviewed in subsection 2.2.2. subsection 2.2.3 summarises the modelling methods used to quantify impacts, categorising them into analytical, probability, regression, and machine learning approaches. The applications of these models reflected in response plans from existing publications, are reviewed in subsection 2.2.4. We provide a discussion of future directions in multi-modal disruption management in subsection 2.2.5, followed by the conclusion of the paper in **??**.

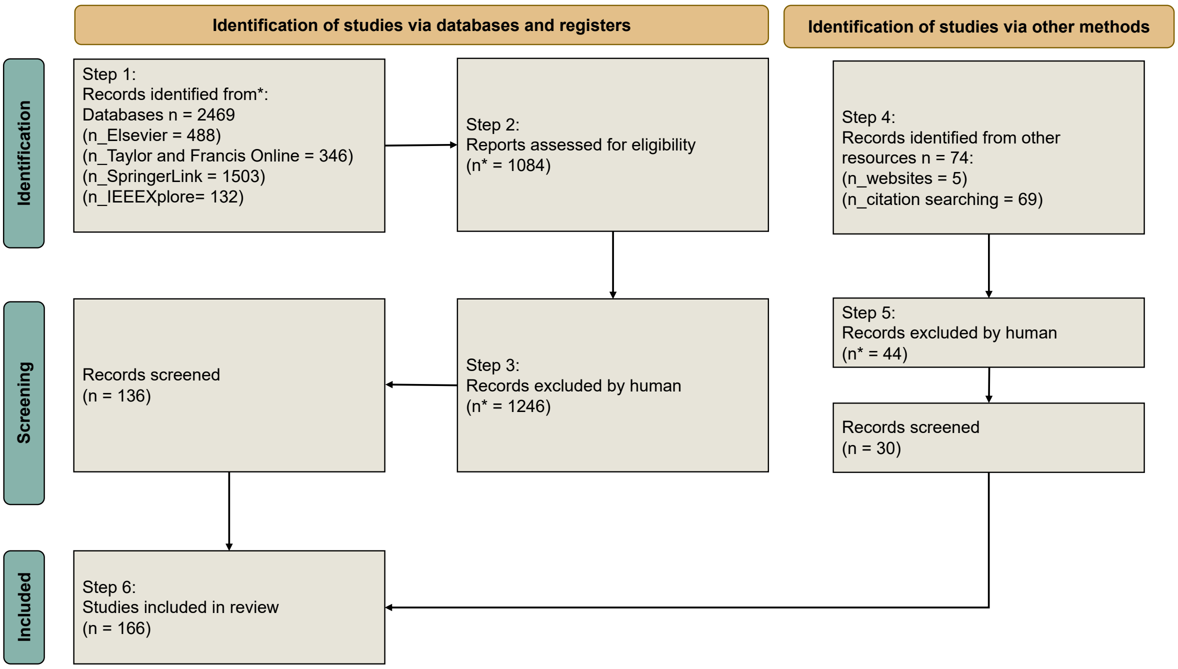


FIGURE 2.4: Flow diagram for systematic review based on the PRISMA approach.

2.2.1 Methodology

Literature review material and PRISMA

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) is the review material selection method that has been applied when organising and analysing literature for this article Page et al., 2021. This method was selected as it is well designed and organised for reporting systematic reviews with objectives, and the clear checklist and flow diagram for review were handy in use. The major processes for reviewing literature are illustrated in Figure 2.4. The first step of this method aims to identify the relevant literature from the databases and other resources, such as websites, organisation publications and citation searching, based on keywords covering multi-modal, PT, disruptions, incidents, modelling, disruption impacts and incident impacts for this research study. The databases accepted for this stage were:

- Elsevier,
- Taylor and Francis Online,
- SpringerLink,
- IEEEXplore.

The database was accessed through the University of Technology Sydney, and the publication date was limited between 2001 and 2024 (the same condition was applied when searching other resources).

From this process, 2469 pieces of literature were collected. After removing ineligible records, 1385 works were retained from the databases, with 1084 works being excluded, as shown in *Step 2* in Figure 2.4. The ineligible records were those that were not open access or to which the author did not have full access. By being filtered by humans according to the literature title, keywords and abstract, 1246 works were further removed, as they were not fully in relation to the purpose of this review, as shown in *Step 3*. After this step, the total number of literature that was ensured to be included in this review article was 136.

Apart from the records obtained from databases, 74 works were found from other resources (5 articles were found on websites and 69 articles were found based on citation searching), as shown in *Step 4 - 5*. In this step, the selected literature was primarily screened by removing duplicates and eligibility, resulting in 30 works being removed. Finally, as shown in *Step 6*, there are 166 works included in this literature review. For each article in the database, the initial review process started with the section abstract and introduction and ended with the section of discussion or conclusion.

Bibliometric analysis

Year According to the literature, the emphasis on multi-modal PT disruption analysis began in 2001, as illustrated in Figure 2.5. There was an initial peak in contributions around 2010, followed by another peak in 2017 and 2020, which observed the highest number of works delivered in a single year. Additionally, there was an exponential increase in the number of publications from 2018 to 2021, followed by a decline to the present.

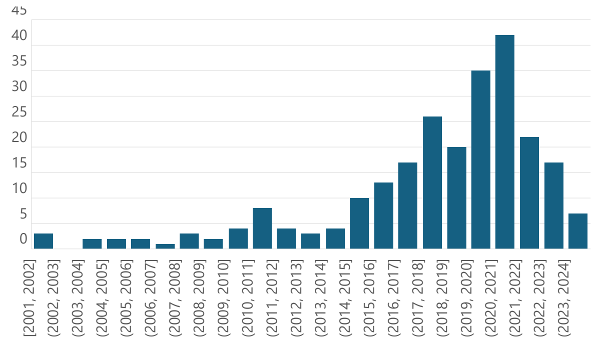


FIGURE 2.5: Histogram of publication years for the period 2001 to 2024..

Authors We present the frequency of author names in the selected database in Figure 2.6. The data shows that the most prolific contributors are Prof. Dr. Oded Cats and Prof. Dr. Francesco Corman, with up to 10 publications primarily focused on PT, capacity, disruption, network vulnerability and subways, and disruption management, agent-based simulation, within-day replanning, computer simulation and information availability, respectively.

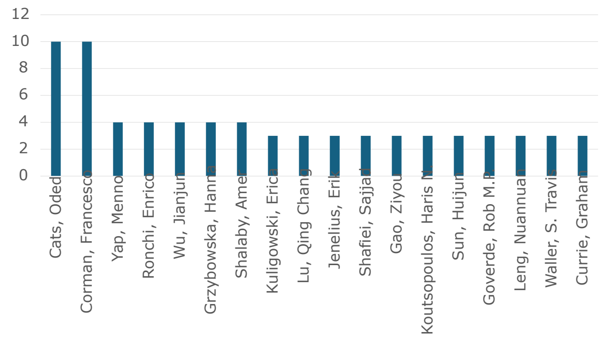


FIGURE 2.6: Count of publications by author for the period 2001 to 2024.

Keywords The keywords were mapped using bibliometric networks in VOSviewer. From all the publications reviewed, the top 50 terms related to multi-modal public transport modelling under disruption were selected based on their co-occurrence in titles, keywords, and abstracts. The terms are further colour-coded by year. As shown in Figure 2.7, terms like vulnerability, disruption, robustness, and complex network related to graph theory-based analysis were more prevalent around 2018. More recent research, around 2022, focuses on COVID-19, transport systems, numerical models and machine learning algorithms.

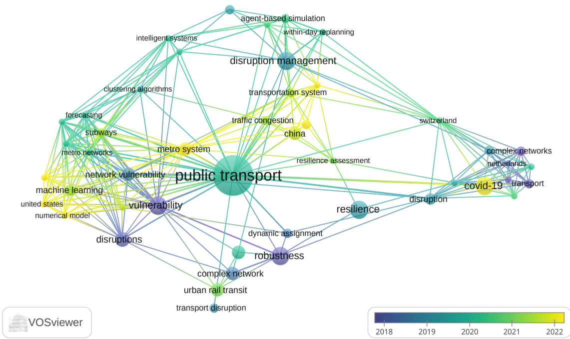


FIGURE 2.7: The network of the hot keywords associated with publications incorporating multi-modal PT under disruption for the period 2001 to 2024.

Modes of transport involved in publications In the field of multi-modal PT modelling under disruption, there are more publications involving only single modes of transport, such as train, metro/subway/underground, bus rapid transit (BRT) or buses. The proportion of publications by transport mode is displayed as Figure 2.8, where disruption navigating the metro/subway/underground network attracts the most attention.

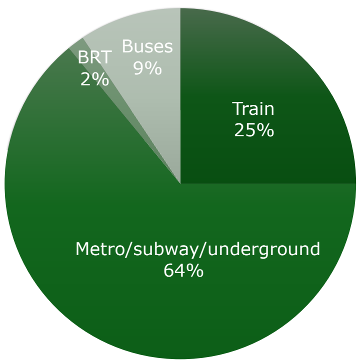


FIGURE 2.8: Overview of transport modes involved in single-mode PT disruption modelling

Among the reviewed publications, 60% included only a single mode of public transport (PT), specifically trains, metro/subway/underground, bus rapid transit (BRT) and buses. The remaining studies involved multiple PT modes, including trains, high-speed trains, trams/light rail, metro/subway/underground, BRT, buses, bikes/shared bikes, shared e-scooters, ferries, airplanes, taxis, or motorbikes. Various combinations of multi-modal PT used in the modelling are shown in Table 2.4. There are 31% research included buses followed by metro/subway/underground, which account for 22%. 16% papers considered trains and trams/light rail.

Reference	Trains	HST	Trams	Metro	BRT	Buses	Bike	E-scooters	Ferries	Airplanes	Taxis	motorbikes
Abdelgawad and Abdulhai, 2010				✓		✓						
Cox, Prager, and Rose, 2011				✓		✓						
Jang et al., 2014			✓		✓							
Cats and Jenelius, 2015				✓		✓						
Neumanna, 2015						✓				✓		
Cats, Yap, and Oort, 2016	✓		✓	✓		✓						
Challender, 2016	✓			✓		✓						
Moylan, Foti, and Skabardonis, 2016	✓		✓			✓						
Cottrill et al., 2017	✓			✓			✓					
D'Andrea and Marcelloni, 2017						✓					✓	✓
Shelat and Cats, 2017	✓		✓									
Yap et al., 2017	✓		✓			✓						
Cats and Jenelius, 2018			✓	✓		✓						
Diab, Feng, and Shalaby, 2018			✓	✓								
Diab and Shalaby, 2018				✓		✓						
Diab and Shalaby, 2018				✓		✓						
Malandri, Fonzone, and Cats, 2018			✓	✓		✓						
Oswald Beiler, Asce, and Miller, 2018	✓						✓					
Nguyen Phuoc et al., 2018	✓		✓			✓						
Yang et al., 2018	✓					✓						

Yap et al., 2018			✓	✓								
Ari Wibowo, Sulistyo, and Mustika, 2019	✓					✓						
Saxena, Hossein Rashidi, and Auld, 2019	✓					✓						
Azolin, Silva, and Pinto, 2020			✓			✓	✓					
Consilvio et al., 2020				✓		✓						
Fang, Jiang, and Fei, 2020			✓			✓					✓	
Jia, Zhang, and Shi, 2020				✓			✓					
Jia, Zhang, and Shi, 2020				✓			✓					
Leng and Corman, 2020b	✓		✓			✓	✓					
Leng and Corman, 2020a			✓	✓		✓	✓					
Marra and Corman, 2020			✓			✓						
Rahimi and Corman, 2020	✓					✓						
Tan et al., 2020				✓		✓						
Aparicio, Arsenio, and Henriques, 2021			✓	✓		✓						
Barbieri et al., 2021	✓		✓	✓		✓	✓			✓		
Cebecauer et al., 2021				✓							✓	
Drabicki, Islam, and Szarata, 2021	✓					✓					✓	
He et al., 2021						✓					✓	
Henry, Furno, and El Faouzi, 2021			✓	✓								
Rahimi et al., 2021	✓					✓						

Yang and An, 2021				✓		✓						
Zhang and Ng, 2021	✓		✓	✓								
Cong et al., 2022				✓		✓						
Liu et al., 2022b				✓		✓						
Mo, Koutsopoulos, and Zhao, 2022	✓					✓						
Yang et al., 2022				✓			✓					
Yap and Cats, 2022			✓	✓								
Yap et al., 2022	✓		✓			✓						
Zhao et al., 2022						✓						
Böcker et al., 2023	✓					✓						
Drabicki et al., 2023			✓			✓						
Feng et al., 2023		✓		✓		✓						
Hasiak, 2023	✓					✓	✓	✓				
Marra and Corman, 2023	✓		✓			✓						
Othman et al., 2023	✓					✓					✓	
Soza-Parra, Tiznado-Aitken, and Muñoz, 2023				✓		✓						
He, Tao, and Sun, 2024				✓		✓						
Zhang et al., 2024a				✓		✓						

TABLE 2.4: Overview of transport modes involved in ,multi-modal PT disruption modelling.

Location of case study involved in publications Figure 2.9 illustrates the distribution of the geographical sources of case studies involved in the literature publications from 2001 to 2024. The most frequently selected case study areas are in the United States, followed by the United Kingdom, the People's Republic of China, Sweden, the Netherlands and Switzerland.

Literature sources Regarding the sources of literature, the top three sources publishing research incorporating multi-modal PT disruption for the period 2001 to 2024 are Transportation Research Record, Safety Science and Accident Analysis & Prevention, as illustrated in Figure 2.10.

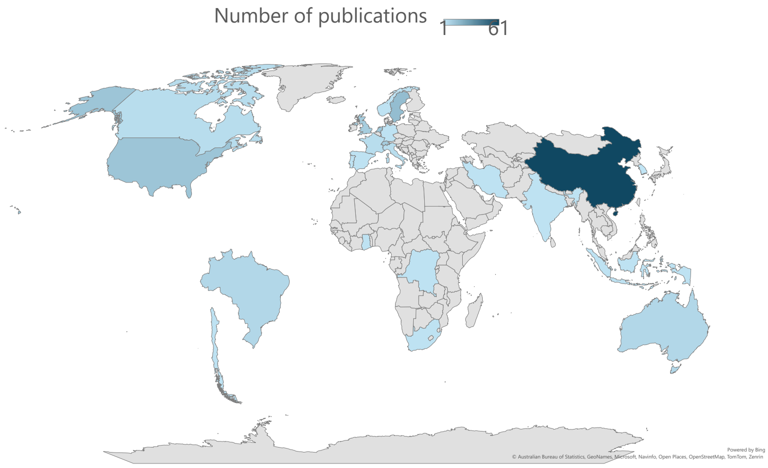


FIGURE 2.9: The distribution of the geographical source of the case studies involved in the literature publications incorporating multi-modal PT under disruption for the period 2001 to 2024.

The trend lines in the chart show the growth of publications over time, with notable rises around certain years: 2012, 2018 and 2021, shown in Figure 2.11. The number of publications by the Transportation Research Record leads almost every year between 2001 and 2024. Notably, the trend for Safety Science indicates a slow start, with minimal publications in the early years, followed by a period of steady growth beginning around 2020. In recent years, however, the trend line has risen sharply, indicating a significant increase in the publication rate incorporating multi-modal PT under disruption.

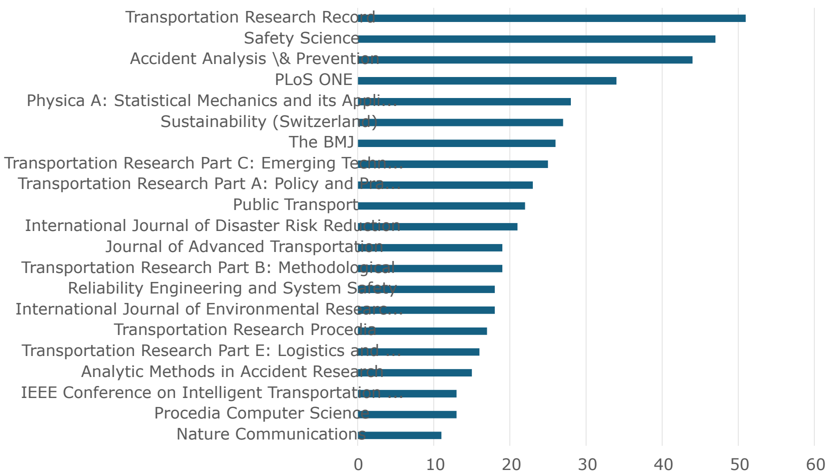


FIGURE 2.10: The total number of literature sources incorporating multi-modal PT disruption for the period 2001 to 2024.

2.2.2 Disruption causes and impact measurement

In studies on traffic disruptions, the most straightforward method to investigate their impact on PT is through direct observation. Depending on the cause of the disruptions, whether from natural disasters or man-made incidents, the investigation covers impacts on travel reliability, including link

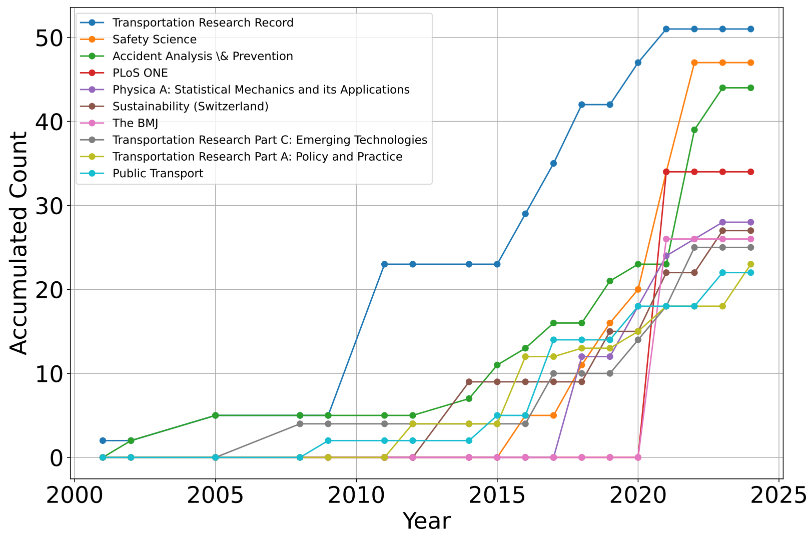


FIGURE 2.11: The cumulative growth of literature sources incorporating multi-modal PT disruption for the period 2001 to 2024.

capacity reduction, node closure, travel demand change and transport performance or supply reduction. When combined with data processing or mathematical modelling approaches, as introduced in subsection 2.2.3, we can precisely quantify the impacts of a given disruption.

Disruption causation

The causes of traffic disruptions discussed in existing publications are summarised in Figure 2.12. Disruptions due to transport service failures receive the most attention, followed by pandemics, and floods, hurricanes and storms. Most research studies found that demand changes due to disruption are the primary impact, followed by reductions in transport performance or supply, impacts on link capacity, and finally, impacts on nodes such as stations or stops. Figure 2.12 summarise the causations based on studies that rely on real disruption case studies; apart from the real-world cases, some research relied on assumed disruptions. Specifically, 29 publications focused on assumed link capacity reduction, 17 on hypothetical node closures and travel demand adjustments, and in 18 publications, the transport performance or supply is manually reduced as negative impacts on the public transport system when modelling.

Natural disaster Flood and sea-level rise

Research studies examining the impact of floods, sea-level rise and heavy rain events on public transport primarily focus on transport infrastructure damage and also investigate the subsequent effects on link or node capacity and vehicle functionality, often reflected by speed reductions. Related research predominantly uses a simulation-based hydrodynamic model when estimating the time or space-dependent water depth based on the historical disaster events and defining the roads as non-functional if the water depth around these roads exceeds a threshold (see Jang et al., 2014; Oswald Beiler, Asce, and Miller, 2018; Zhu et al., 2022. After investigating the floods between January and November 2021, Yang et al., 2018 identified the impacts of a tropical cyclone in Hainan, China, on bus stations, railways and highways damage, which caused connectivity, efficiency reduction and travel

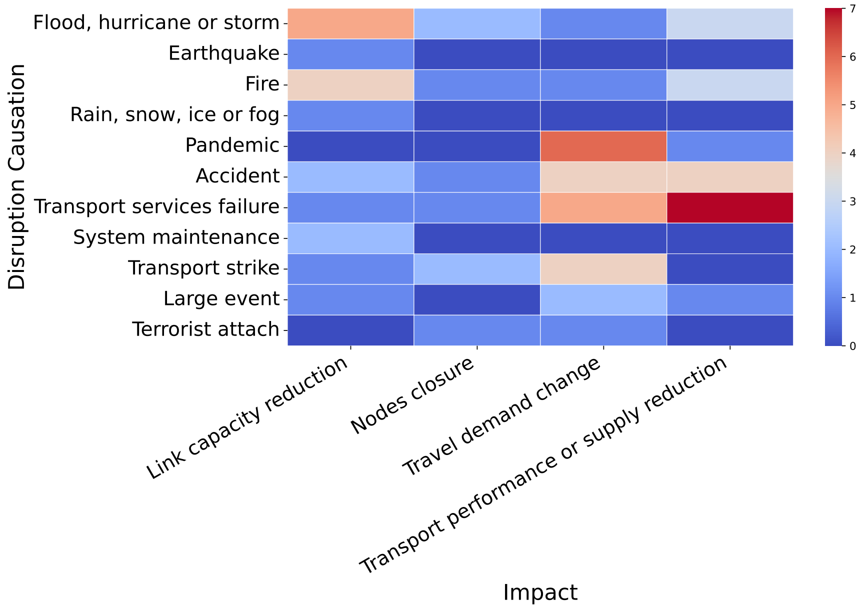


FIGURE 2.12: The heatmap of the publication regarding the causation associated with multi-modal PT disruptions.

delays. Zhu et al., 2019 captured the capacity loss of links due to Hurricane Sandy in New York according to local taxi trajectory data.

Most research on the impact of floods has primarily focused on cars, as highlighted in review papers by Hammond et al., 2015; Munawar, Hammad, and Waller, 2021; Bhuiyan et al., 2021). Pregnolato et al., 2017 reviewed and proposed a depth-disruption function to model the impact of the depth of standing water on road vehicle speed and traffic flow.

Earthquake

When estimating the impact of earthquake/seismic events on road networks, the fragility function was used in most of the studies, where the impact of the seismic intensities on road networks can be qualified. Similar to the studies on the impacts of flood, historical data regarding the ground motion of a seismic event against location or time was mostly used to determine the seismic intensities and the threshold towards railway network degrading or damage (see Ansari, Rao, and Jain, 2024).

Fire

Regarding the assessment of disruption impacts due to fire, Ronchi et al., 2019; Wahlqvist et al., 2021; Cova et al., 2021; Lovreglio et al., 2020 investigated the impact of bushfires or wildfires, while Ren et al., 2019; Bjelland et al., 2021 examined the impact of tunnel fires on transit networks. Additionally, Wetterberg, Ronchi, and Wahlqvist, 2021 focused on the impact of wildfire smoke. Readers can find more references regarding the impact of fire events on road networks from the review article of Ronchi et al., 2019.

Climate-related

Many studies have utilised surveys and observational data to examine the impact of snow, ice, heavy rain, or fog events on transport networks. Jang et al., 2014 created a snowfall map using snow depth, sunny spots and shaded location data corresponding to the tram network to capture the time and space impact of snow cover on the tracks, affecting tram link accessibility. Yasanthi and Regina, 2018 indicates that climate-related or weather-related events most commonly result in reduced travel speeds due to traffic congestion caused by low visibility or road closures for safety reasons. Inatsu, Tanji, and Sato, 2020 explored the correlation between blowing snow levels and expressway closures, while Elhenawy, Rakha, and Ashqar, 2021 analysed the effects of both light and heavy rain on travel speed. Page-Tan and Aldrich, 2024 specifically examined the impact of the Snowmageddon event in Boston. Previous studies on weather-related transport disruptions have been summarised in the review by Jarmuz and Chmiel, 2020.

Pandemic

The COVID-19 pandemic is creating disruptive changes in urban mobility, especially PT. Travel demand was largely reduced (see Aparicio, Arsenio, and Henriques, 2021; Soza-Parra, Tiznado-Aitken, and Muñoz, 2023; Garcia-Arteaga and Lotero, 2023) or mode choice (see Du et al., 2024) was adjusted due to the fear of the pandemic; PT supplement was reduced due to the operational staff reduction because of the pandemic (see Soza-Parra, Tiznado-Aitken, and Muñoz, 2023). Böcker et al., 2023 discusses the long-term impact of the pandemic on travel demand and travel behaviour change based on a survey result. Feng et al., 2023 proposed air change rates in buses, subways, and high-speed trains based on measured CO2 concentrations and passenger behaviours. Using these rates, the author suggested safe times and spaces in various public transport vehicles, which differed significantly from the patterns used in normal situations.

Man-made incidents For man-made incidents, historical incident logs are the primary sources of information for investigating the causes of disruptions. Extensive research has utilised data-driven algorithms (regression, probability or Machine Learning (ML) algorithms) to understand the causation of disruptions for cars (see Li, Pereira, and Ben-Akiva, 2018; Yang, Zhang, and Feng, 2022). Yap and Cats, 2021 applied logistic regression and a Multilayer Perceptron (MLP) classifier when predicting disruption based on historical incident log data. This author also used a supervised learning approach to quantify passenger delays as the impact of PT disruption. However, to analyse the temporal-spatial impact on public transport, most research relies on data analysis, including:

- time or location-dependent passenger volume count or flow data (see Zhao et al., 2022; Kopsidas and Kepaptsoglou, 2022; Moylan, Foti, and Skabardonis, 2016; Jia, Zhang, and Shi, 2020; Marra and Corman, 2020; Cebecauer et al., 2021; Henry, Furno, and El Faouzi, 2021; Cong et al., 2022; Mo, Koutsopoulos, and Zhao, 2022; Yang et al., 2022; Yap and Cats, 2022; Wen et al., 2018; Liu, Zhu, and Wang, 2021; Zhao et al., 2021; Zhao et al., 2023),
- travel behaviour (mode or route choice) (see Cottrill et al., 2017; Nguyen Phuoc et al., 2018; Saxena, Hossein Rashidi, and Auld, 2019; Jia, Zhang, and Shi, 2020; Cong et al., 2022; Mo, Koutsopoulos, and Zhao, 2022; Yang et al., 2022; Böcker et al., 2023; He, Tao, and Sun, 2024; Eltved et al., 2021),
- travel speed information (see Moylan, Foti, and Skabardonis, 2016; Diab and Shalaby, 2018; Othman et al., 2023; Shafiei et al., 2021a),
- travel time data (see Zhao et al., 2022; Kopsidas and Kepaptsoglou, 2022; Jia, Zhang, and Shi, 2020; Cong et al., 2022; Yap and Cats, 2021; Müller, Leich, and Nagel, 2020),

- trip trajectory or vehicle GPS data (see D'Andrea and Marcelloni, 2017; Marra and Corman, 2020),
- topological features, such as link criticality (see Cats, Yap, and Oort, 2016).

According to the literature, the reason for man-made accidents that disrupt traffic in PT can include crashes between metro vehicles (see Mo, Koutsopoulos, and Zhao, 2022). According to 167 metro incident reports, Lu, 2018 found that the most frequent occurrences are train breakdowns, signal control failures, and door malfunctions. The impacts of these operational incidents on passengers include delayed travel, unmet travel demand and accidents caused by overcrowding. After analysing 408 metro incidents, Lu, Ma, and Xing, 2021 investigated Shanghai metro disruption data from 2013 to 2016, comprising 583 samples, and found that train failures accounted for 45% of the total incidents, followed by signal failures at 20%. Liu, Zhu, and Wang, 2021 indicated that 48.7% disruption was directly caused by signal failure, followed by train failure (23.0%), and 12.7% disruption was due to passengers falling into the track. For transport services failure or mode of transport shut down, the further reason could be a power, signal or telecommunication failure (see Liu, Zhu, and Wang, 2021; Zhao et al., 2021). Yang, Dong, and Guo, 2023 analysed incident log data and found that metro disruption was majorly due to train failure (29%) followed by a signal failure (22%), and more disruptions occurred during the morning peak due to the frequency of services.

The disruption reason of large events includes in related literature majorly investigate the demand change and PT supplement change due to events like the 2014 Commonwealth Games (Cottrill et al., 2017) and 2017 Lyon's "Festival of Lights" (Henry, Furno, and El Faouzi, 2021).

Other causes for traffic disruption include system maintenance, transport strikes or terrorist attacks, which have been documented in the literature, as summarised in Table 2.5.

In addition to analyses based on real events, numerous research studies have focused on the impact of assumed disruptions. These studies primarily relied on simulation models and applied assumed impacts, such as link capacity reduction, node closure, travel demand or behaviour changes, transport performance, or supply reduction, during the scenario analysis stage. By comparing traffic states before and after the inclusion of the disruption, these models can visualise the impact of the disruption properly.

Disruption causes	References			
	Link capacity reduction	Nodes closure	Travel demand change	Transport performance or supply reduction
Natural disaster				
Flood, hurricane or storm	Jang et al., 2014; Challender, 2016; Oswald Beiler, Asce, and Miller, 2018; Zhu et al., 2019; Pregnolato et al., 2017	Challender, 2016; Yang, Dong, and Guo, 2023	Zhu et al., 2022	Yang et al., 2018; Zhu et al., 2022; Pregnolato et al., 2017
Earthquake	Ansari, Rao, and Jain, 2024			
Fire	Ronchi et al., 2019; Wahlqvist et al., 2021; Ren et al., 2019; Bjelland et al., 2021	Ronchi et al., 2019	Cova et al., 2021	Wahlqvist et al., 2021; Cova et al., 2021; Ren et al., 2019
Rain, snow, ice or fog	Jang et al., 2014			

Pandemic			Aparicio, Arsenio, and Henriques, 2021; Böcker et al., 2023; Soza-Parra, Tiznado-Aitken, and Muñoz, 2023; Garcia-Arteaga and Lotero, 2023; Du et al., 2024; Borowski et al., 2023	Soza-Parra, Tiznado-Aitken, and Muñoz, 2023
Assumption	Papangelis et al., 2013	Zhang and Ng, 2021; Yang and An, 2021	Yang and An, 2021; Böcker et al., 2023	Yang and An, 2021
Man-made incident				
Accident	Zhao et al., 2022; Mo, Koutsopoulos, and Zhao, 2022	Mo, Koutsopoulos, and Zhao, 2022	Zhao et al., 2022; Mo, Koutsopoulos, and Zhao, 2022; Sun et al., 2016; Liu, Ma, and Koutsopoulos, 2021	Othman et al., 2023; Yap and Cats, 2021; Liu, Ma, and Koutsopoulos, 2021; Yap and Cats, 2019
Transport services failure	Cats, Yap, and Oort, 2016	Jia, Zhang, and Shi, 2020	Saxena, Hossein Rashidi, and Auld, 2019; Liu, Zhu, and Wang, 2021; Zhao et al., 2023; Zhu et al., 2017	Diab and Shalaby, 2018; Zhao et al., 2021; Lu, 2018; Lu, Ma, and Xing, 2021; Khadilkar, 2017; Mahdavi, Bhouri, and Scemama, 2020; Zhang et al., 2022
System maintenance	Yap and Cats, 2022; Eltved et al., 2021			
Transport strike	Moylan, Foti, and Skabardonis, 2016	Yang et al., 2022; He, Tao, and Sun, 2024	Moylan, Foti, and Skabardonis, 2016; Nguyen Phuoc et al., 2018; Yang et al., 2022; He, Tao, and Sun, 2024	
Large event	Cottrill et al., 2017		Cottrill et al., 2017; Henry, Furno, and El Faouzi, 2021	Cottrill et al., 2017
Terrorist attack		Tokuda et al., 2006	Cox, Prager, and Rose, 2011	

Assumption	Cats and Jenelius, 2015; Shelat and Cats, 2017; Yap et al., 2017; Malandri, Fonzone, and Cats, 2018; Consilvio et al., 2020; Leng and Corman, 2020b; Rahimi and Corman, 2020; Müller, Leich, and Nagel, 2020; Wang et al., 2021b; Liang et al., 2019; Kepaptsoglou and Karlaftis, 2009; Jin, Teo, and Odoni, 2016; Zhu and Goverde, 2019; Lu and Lin, 2019; Xu and Liang, 2020; Leng, Liao, and Corman, 2020; Tang et al., 2021b; Malandri et al., 2021; Shao and Song, 2022; Tan et al., 2023; Massobrio and Cats, 2024; Chen et al., 2024; Zhang et al., 2024b	Kopsidas and Kepaptsoglou, 2022; Cebecauer et al., 2021; Drabicki, Islam, and Szarata, 2021; Cong et al., 2022; Jin, Teo, and Odoni, 2016; Tang et al., 2021b; Zhang et al., 2023; Shi et al., 2019; Xing et al., 2017; Zhang, Wang, and Wang, 2018; Yin et al., 2018; Zhang et al., 2018; Wang, Yuan, and Yin, 2019	Kopsidas and Kepaptsoglou, 2022; Cebecauer et al., 2021; Drabicki, Islam, and Szarata, 2021; Cong et al., 2022; Jin, Teo, and Odoni, 2016; Tang et al., 2021b; Zhang et al., 2023; Shi et al., 2019; Xing et al., 2017; Zhang, Wang, and Wang, 2018; Yin et al., 2018; Zhang et al., 2018; Wang, Yuan, and Yin, 2019	Abdelgawad and Abdulhai, 2010; Ari Wibowo, Sulistyo, and Mustika, 2019; Azolin, Silva, and Pinto, 2020; Marra and Corman, 2020; Tan et al., 2020; Marra and Corman, 2023; Othman et al., 2023; Liu, Ma, and Koutsopoulos, 2021; Zhu and Goverde, 2019; Leng, Liao, and Corman, 2020; Srikukenthiran and Shalaby, 2017; Placido, 2017; Dollevoet et al., 2018; Dekker and Panja, 2021; Corman, 2020; Miristice et al., 2023

TABLE 2.5: The disruption causes and their impacts investigated in the literature.

Disruption measurement

Traffic disruptions can lead to accidents and hazardous conditions that result in inefficiencies in the transportation network and economic repercussions, including lost productivity, increased operational costs for businesses and higher transportation costs. According to the existing publications, the typical impacts on travel reliability are majorly reflected by:

- roads or links fully or partially closed or link capacity reduction,
- PT stations or trip production or attraction nodes fully or partially closed,
- change in travel demand or travel behaviour,
- transport performance or supply reduction.

In order to better compare the publications on disruption impacts, we summarised various impacts in Table 2.5, where transport performance reduction refers to the impact that reduces the number of PT runs or trips or PT vehicle travel speed reduction.

In terms of acquiring the impact on road networks, including the degrading of road links, network nodes, demand and travel behaviour as well as the transport supply, assorted data processing approaches are selected for measuring the impact of disasters on transport networks, the methods include:

- simulation-based experimental approaches,
- survey,
- lab-based experimental approaches,

- detected,
- field observations.

Various measurement methods capture multiple types of disruption impacts. As summarised in Figure 2.13, simulation-based experimental approaches are commonly used in existing research studies. By using simulation-based approaches, roads or links that are fully or partially closed or link capacity reduction and changes in travel demand or travel behaviour are mostly obtained as negative impacts on PT networks. Data collected from detectors and field observations help capture the impact on transport performance decay or supply reduction. Surveys are a robust method for collecting data on demand changes in response to transport disruptions.

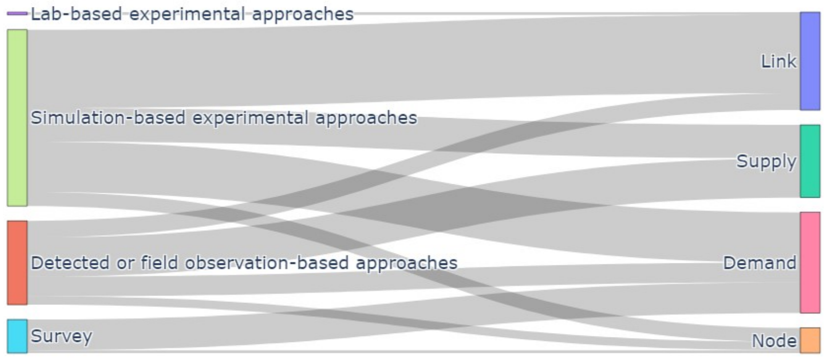


FIGURE 2.13: The Sankey Diagram of the measurement methods associated with multi-modal PT disruption impacts.

Simulation-based experimental approaches The impact analysis using the transport simulation-based models tends to simulate the change in network performance or the change of movement in a post-disruption situation. The changes are treated as the impact of disruptions. The disruptions are often described by hypothesis attacks that refer to link failure, PT station or node failure, change of demand or travel behaviour or transport supply reduction, as summarised in Table 2.5. The use of this experimental approach is mainly due to the lack of historical disruption or incident log data. Since analytical results highly rely on the quality of both the transport modelling and the disruption modelling, the way to prove the accuracy of modellings always challenges the researchers who apply such an experiments-based approach. This type of modelling method can be categorised as discrete or continuous by time or space and can be applied at the macroscopic (see Yang and An, 2021; Yang et al., 2018; Tan et al., 2020; Henry, Furno, and El Faouzi, 2021; Yap and Cats, 2022; Liu, Zhu, and Wang, 2021; Zhao et al., 2021; Zhao et al., 2023; Liang et al., 2019), mesoscopic (combined of macroscopic and microscopic) (see Abdelgawad and Abdulhai, 2010; Sun et al., 2016), or microscopic (see Leng and Corman, 2020b; Yap et al., 2022; Othman et al., 2023; Müller, Leich, and Nagel, 2020; Liu, Ma, and Koutsopoulos, 2021; Wang et al., 2021b; Wang, Yuan, and Yin, 2019) levels.

Those research studies that rely on a discrete event model, such as D'Andrea and Marcelloni, 2017; Consilvio et al., 2020; Tan et al., 2020; Drabicki, Islam, and Szarata, 2021; Soza-Parra, Tiznado-Aitken, and Muñoz, 2023; Zhu and Goverde, 2019; Chen et al., 2024; Yin et al., 2018; Corman, 2020, tend to simulate transport systems where changes occur at distinct points in time. These changes, or events, are discrete and represent significant occurrences that alter the state of the system. A continuous model, regardless of the event or state change at a certain time point, mimics the continual change of system states following the natural time progression. The state of the system is described by differential equations that are solved over continuous time intervals.

In macroscopic modelling methods, variables such as average speed, flow, and density are primarily used to represent traffic conditions. Traffic is often modelled as a continuous flow rather than focusing on individual vehicles or users. Conversely, microscopic modelling methods concentrate on individual units, such as vehicles or users, and their specific trips. Mesoscopic modelling serves as a hybrid approach, combining elements of both macroscopic and microscopic modelling. It may focus on individual vehicles but aggregate their behaviour over larger segments. Among simulation models used in existing publications, the microscopic modelling method is the most popular. There is much mature commercial software that is able to simulate networks relatively accurately. The commonly used simulation software mentioned in the literature is demonstrated inFigure 2.14.

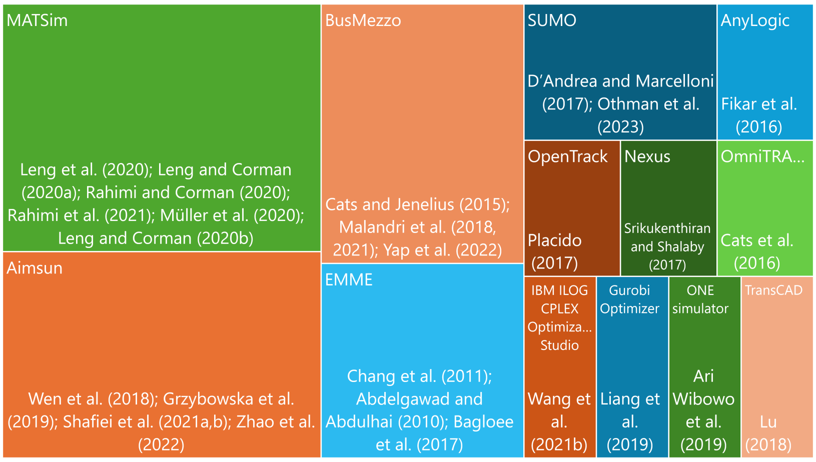


FIGURE 2.14: The simulation software mentioned in the literature associated with multi-modal PT disruption impacts.

A trip distribution simulation model can also be programmed by using C++ (see Wang, Jin, and Sun, 2022), Python (see Yin et al., 2022) or Matlab (see Fang, Jiang, and Fei, 2020; Mahdavi, Bhouri, and Scemama, 2020; Lu and Lin, 2019), based on the programmed testbed, a remove of the node or blocked link can be applied to capture the impact on trip distribution or travel delay.

Using the simulation model, scenarios before and after a disruption can be simulated. By comparing these changes, it is possible to uncover the impact of such disruptions. Malandri, Fonzone, and Cats, 2018 compared the change in Passenger Volume Over Capacity (VOC) ratio before and after an unplanned service segment disruption to analyse its evolving impact. Zhao et al., 2022 relied on a model developed using commercial software (Aimsun) while providing a framework for testing the impacts of various road section blockages on the changing traffic states, including delay, flow and density over the entire study area. Three types of experiments were considered to understand the impact of disruption duration, scale, and the reaction of PT users under a disruption. Wetterberg, Ronchi, and Wahlqvist, 2021 applied the VR simulation developed by the game engine Unity3D when capturing the impact of fire smoke on driving behaviour reflected in vehicle speed. The agent-based simulator named BusMezzo was adopted by Malandri et al., 2021 when modelling trajectory for individual travellers and effects of traffic disruption reflected by link capacity decrease.

Survey Some research studies investigate the impact through transport surveys, where assumed disasters and their potential impacts, such as post-accident scenarios, road closures, or PT delays, are included. This approach allows interviewees to demonstrate their decision-making in response to specific impacts. By doing so, we can estimate changes in demand or travel behaviour resulting from the disruption for relevant research. Generally, the data collected by a transport survey targets individual opinions, which helps obtain information for building microscopic models. In existing research studies, surveys have been used to collect performance data in areas including Australia, China, France, Sweden, the United Kingdom and the United States, as shown in Figure 2.15.

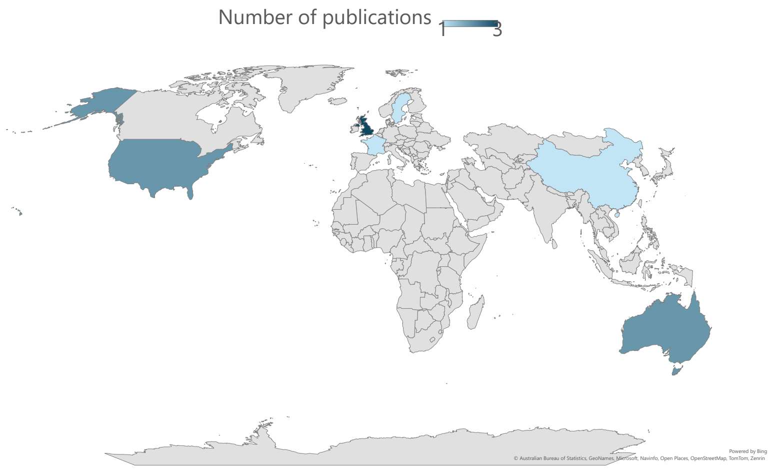


FIGURE 2.15: The distribution of the geographical source where surveys are conducted for obtaining demand change.

Cox, Prager, and Rose, 2011 studied the decrease in journeys following subway and bus bombings in London, particularly due to partial or full station closures. Approximately 22.7 million passenger journeys were reduced, with about 9 million attributed to station closures and the remaining 13.6 million attributed to fear factors.

Papangelis et al., 2013 conducted 52 interviews in rural areas of Scotland and England to explore passenger preferences regarding route and mode choice during disruptions, aiding demand estimation under such conditions. The disruptions considered in the survey included accidents reducing link capacity, bus delays, heavy winds reducing link capacity, and congestion.

Cottrill et al., 2017 investigated the impact of the 2014 Commonwealth Games in Glasgow on transport control and interviewed and explored the role of social media, such as Twitter, in managing travel behaviour.

Saxena, Hossein Rashidi, and Auld, 2019 conducted interviews in Chicago to understand attitudes toward alternatives in response to delayed or cancelled transit services. The alternatives considered included waiting for the same transit service, ride-sharing (e.g., Uber or Lyft), taking a taxi, getting picked up by family or friends, changing the destination, and cancelling the trip.

Nguyen Phuoc et al., 2018 investigated the likelihood of mode shifts resulting from service cancellations during a strike in Melbourne, Australia, through survey-based research. The study found that during such disruptions, 39.4% of train users would shift to driving a car, approximately 40%

would switch to other public transport modes (tram and bus), and 6.6% indicated they would cancel their trips altogether.

Auld et al., 2020 ran a survey to obtain responses about the change in travel behaviour against service disruptions in the Chicago metropolitan area. The results showed that travellers intend to maintain using the transport mode, either waiting for service recovery or bridging shuttles (the alternative shuttles provided for journey recovery); fewer travellers would turn to another available transport mode, such as taxis.

Hasiak, 2023 analysed changes in student travel behaviour in France due to the pandemic, using data from 13 household travel surveys.

Böcker et al., 2023 conducted interviews with the public in three Nordic cities, Oslo, Bergen and Stockholm, comparing surveys from 2020 (N=2,718) to 2018 (N=3,523). The study found a significant mode shift from public transport to private cars post-pandemic, leading to a demand adjustment for future transport modelling.

Du et al., 2024 conducted a survey of 1,045 residents in NSW, Australia, between October 2021 and May 2022 to understand changes in travel behaviours due to the pandemic. The results identified factors influencing travellers' choices regarding travel mode, travel purpose, and acceptance of emerging mobilities such as on-demand transport, autonomous vehicles and drones.

He, Tao, and Sun, 2024 collected questionnaire surveys from late June to early July 2020 to investigate the disruptive effects of Hong Kong social movements on a transit system.

Lab-based experimental approaches The lab-based experimental approaches literally refer to the method of conducting physical experiments aiming to build the disruptions in the laboratory and analyse the actual impacts. Due to the limited lab capacity, experiments can only capture the impact of an individual vehicle due to an incident, which means that this method can hardly be applied to uncover the impact on transport networks. Liu et al., 2022a focused on the impact between car and e-bike crashes, and the experiments-based impact data was fitted to a multiple linear regression model.

Detected or field observation-based approaches **Detectors**

Under this category, and besides the infrastructure-based detectors, the sensor installed inside a vehicle should also be contained; such techniques are employed to capture travel speed, including acceleration or deceleration, location and travel or driving environment, including vehicle temperature, on-road obstacles or potential transport hazards Rettore et al., 2019. The typical technologies used as in-vehicle sensors include global positioning system (GPS) and Bluetooth (see SUWANNO et al., 2021), where the information captured can be utilised to illustrate travel trajectories that image the travel reliability from the individual vehicle point of view. Zhang et al., 2022 used a Gaussian Mixture Model (GMM) on headway deviations captured by AVL data to identify clusters of abnormal headways. The minimum probability of an observation falling into the abnormal cluster was used to classify it as disrupted.

The General Transit Feed Specification (GTFS) is a common standard format for PT schedules and associated geographic information. GTFS is divided into two main parts: GTFS Schedule and GTFS Real-time. GTFS Schedule includes information about routes, schedules, fares, and geographic transit details. GTFS Real-time provides trip updates, vehicle positions and service alerts. This data is widely used in PT modelling for disruption identification and impact measurement reflected by demand; see Aparicio, Arsenio, and Henriques, 2021; Soza-Parra, Tiznado-Aitken, and Muñoz, 2023; Zhao et al., 2023. Research conducted by He et al., 2021; Soza-Parra, Tiznado-Aitken, and Muñoz, 2023; Yap and Cats, 2021; Zhang et al., 2022; Massobrio and Cats, 2024; Srikukenthiran and Shalaby, 2017 focused on the change in PT performance reflected by headway or PT reliability.

Regarding PT usage and individual trip trajectories, smartcard data is essential as it captures passenger volumes and travel behaviour. These datasets illustrate the trajectories of PT users and the demand for specific transport modes. This information helps infer the impact of disruptions on passenger flow and network behaviour. Research by Zhao et al., 2023; Sun et al., 2016; Liu, Ma, and Koutsopoulos, 2021 focused on changes in passenger flow, while Yap et al., 2017; Mo, Koutsopoulos, and Zhao, 2022; Eltved et al., 2021 examined travel behaviour and decision-making related to mode changes. Sun et al., 2016 investigated the change of station choice based on smartcard data and found that under two selected incidents, the detour probability for passengers who chose the neighbouring station was about 13.5% and 11.5%, respectively.

Rusmawati and Rismala, 2016 collected text data from the official Twitter account of the commuter line to identify disruption categories, dates, days, times, locations, number of affected trains, directions (bound for), and positions where the disruptions occurred.

Field observation

Field observation includes the actual in-person examination of the impacted area and recording the features of the area to construct an impact description or record reported to the authorities (see Challender, 2016; Ansari, Rao, and Jain, 2024; Lu, Ma, and Xing, 2021; Tokuda et al., 2006; Zhang et al., 2016).

The data known as an incident log is one typical authority-provided data, where it includes not only the details related to the incident, such as start and end time, duration, location or incident types or reasons but also the scale of the incident embodied by the number of impacted section lanes. Louie, Shalaby, and Habib, 2017 investigated the probability of subway delay based on the incident log data from Toronto's subway system for the year 2013. Lu, 2018 investigated 167 operational incidents that were reported in the metro system in Shanghai, China. Diab, Feng, and Shalaby, 2018 investigated 924 streetcar and 144 subway incidents in Toronto, causing daily delays of 216.7 and 34.4 minutes, respectively. Wen et al., 2018; Grigorev et al., 2022b trained an ML-based incident prediction model based on historical incident log data. Liu, Zhu, and Wang, 2021 investigated 408 incident records in the metro in order to capture the major causes and impacts of metro disruption on passenger flow change. Mo, Koutsopoulos, and Zhao, 2022 measured the change of mode and route choice of metro passengers under incident based on the record in incident log data. Yin et al., 2022 gathered field data that recorded all the disruption events in the Beijing Metro network from 2013 to 2018 and proposed a probability graph-based Bayesian Network (BN) model to predict the disruption.

2.2.3 Disruption modelling methodology

Traffic disruptions impact the performance of road networks, which directly disturbs mobility, city productivity and the growth of the economy. Such consequences have attracted increasing attention and brought learning from observation of previous traffic disruptions, in particular, to uncover the details of transport networks and passenger movement subjected to the traffic disruptions. Apart from observation data, the accuracy of modelling disruption heavily depends on the development of the transport network representation and the transport network modelling approaches, as the estimated impact of a modelled disruption is calculated according to the mathematical relationship between entities involving modelling a transport network. In previous section 2.2.2, we summarise the common methods of collecting impact data; in this section, we extend the topic and provide a summary of modelling methods that are commonly employed for impact quantification.

Based on the literature, the commonly used disruption modelling approaches include analytical approaches, probability-based approaches, regression analysis approaches and ML-based approaches. The applications of the mentioned network and disruption modelling approaches are summarised in Figure 2.16. Based on the figure, a majority of research studies rely on a simulation method, followed by initial stage regression and ML algorithms.

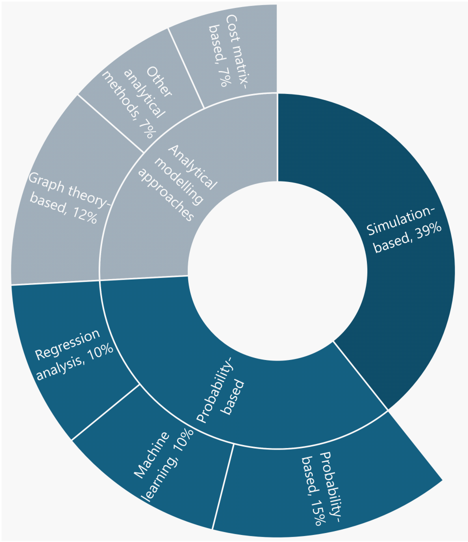


FIGURE 2.16: The type of disruption modelling methods involved in the existing literature.

Analytical modelling approaches

The analytical-based modelling method refers to the methods that rely on scenario experiments, where various disruptions are applied to the network with a given network model, and the change in flow or traffic mobility on the network is the impact of such disruption. The disruptions are hypothesised in time, location and duration according to incident log data or experiences, where these three elements, time, location and duration, are the disruption features that define a type of disruption. The hypothesis demonstrates the impacts described in subsection 2.2.2. Therefore, with the change of disruption time or location or the increase in duration, we are able to analyse the relationship between different types of disruptions and their impacts on modelled networks. The analytical approach suits any type of network model, even if it is not a comprehensive simulation model.

Cost matrix

Modelling a disruption based on a travel cost matrix is a typical example. The cost matrix is filled with variables that change along with the adjustment of other variables. This approach offers a means to observe the discrete changes in the transport system performance before and after a traffic disruption or before and after a disruption intervention management. The cost matrix method does not require establishing the network representation or flow representation; all it needs is the costs at nodes or links according to the logical relationship within a transport system. Typical elements of the cost estimation include travel time, travel distance, the charge for travelling, delay, transfer time, transfer distance, walking time, walking distance and so on. Bell et al., 2017 found a capacity-weighted adjacency matrix representation of a graph by considering the Laplacian matrix. The use of the Laplacian matrix was able to abstract the networks' structure properties while exposing the network performance on node connectivity. A hypothesised reduction of link capacity captured the road section vulnerability. Khadilkar, 2017 developed a delay probability vector based on observed delay data. Azolin, Silva, and Pinto, 2020 discussed the impact of disruptions on travel demand and bus service reliability by assuming the impact distance for each bus route using a region-based matrix. Aparicio, Arsenio, and Henriques, 2021 compared the demand matrix before and after a disruption to capture the change in travel patterns. Garcia-Arteaga and Lotero, 2023 proposed a directed graph to represent travel patterns before and after the pandemic lockdown for a Bus Rapid Transit system in Bogotá. Zhu et al., 2022 converted the nationwide railway network and flooding depth into cost matrix layers and matched these layers to capture the impact of floods using a failure threshold based on flood depth.

Graph theory

The cost matrix is an application in graph theory to demonstrate the travel cost against locations. More detailed analytical transport models utilising graph theories at their cores tend to include the entire network. The graph theory-based modelling approach abstracts the topology of a network and makes use of the topological properties and indicators to represent the performance of the transport system. In a traditional graph theory-based model, transport networks are described by nodes and edges, where nodes are used to represent intersections, residential centres, business centres, PT stations or terminals for freight transport, while edges are used to represent the connections between nodes. The usage and capacity of the transport system are represented by weights on nodes or links. Since the components of a theory-based model are nodes and links, the removal of the nodes or links can simply simulate the disruptions on those nodes or links (see Yang et al., 2018; Massobrio and Cats, 2024; Shi et al., 2019; Xing et al., 2017; Zhang et al., 2018; Jiang, Lu, and Peng, 2018). Most of the previous transport network statistical analyses that aimed to identify the vulnerable links or nodes also highly relied on this theory-based modelling approach. Apart from node or link failure (removal), a reduction of link capacity can be one impact of disruption. Shelat and Cats, 2017 proposed that a disrupted link results in a reduction of link load due to infeasible interchanges and simulated the disruption scenarios by using an L-space graph.

Among the studies of disruption impact considering graph theory, some models are built dynamically to display the cascade effect of a disruption in the network. This concept depends on the node capacity, where a failure of a node results in a flow re-assign to its adjacent nodes; the extra flow would further result in excess capacity at the adjacency, which, in this way, the node failure is spread in the network. This model is known as the load-capacity model: the failure of a node resulted in an increase in load (see Chen et al., 2021). In research studies of Cao and Cao, 2020 and Wang et al., 2021a, the node failure was simulated by the removal of nodes either randomly or intentionally to demonstrate the random attack and malicious network attacks. A similar modelling method was adopted by Tang et al., 2021a; Xing et al., 2017; Zhang, Wang, and Wang, 2018 and Zhang and Ng, 2021, but they only simulated the random attack on a network. In a study of Cumelles, Lordan, and Sallan, 2021, the failed node was set by reducing the node capacity to zero.

The impact of transfer capacity between rail and bus services on cascade failure in a network was estimated by Liu et al., 2022b. In research of Liu et al., 2022b, the failure of nodes was simulated by limiting the access to rails or buses, while the cascade failure was triggered by passengers transferring to another means of transport in the multi-modal PT network. Such cascade failure of disruption by relative time is fundamental to dynamic impact modelling that is absolute time-dependent. Rather than completely removing a node from the network, reducing a node's capacity also negatively affects network robustness. This was demonstrated by Yang and An, 2021, who investigated the impact of node capacity on network failure. Corman, 2020 simulated a train delay spread within the network. Yang, Dong, and Guo, 2023 modelled the flood-triggered cascading failure of a metro system in Zhengzhou, China. The model used a graph where each node was weighted by the probability of impact, identifying the vulnerable nodes in the metro network against flooding. Zhang et al., 2023 investigated the bus network vulnerability towards stop failures.

The reduction of network links' capacity is also a reason for traffic disruption. Nguyen Phuoc et al., 2018 applied the four-step transport model within the Cube software platform to compare trip assignments before and after service cancellations due to disruptions, capturing how these disruptions influenced changes in travel patterns. In a study of Ma et al., 2021, a train delay model against time and space was created by considering the max-plus algebra theory, where train disruptions were modelled by a hypothesised reduction of transport capacity or train speed.

On the other hand, the failure of the network is also in relation to unexpected flow.Shen, Ren, and Ran, 2019 played with extra flow as the perturbation in the cascading failure model to mimic and investigate the cascading failure spread in a metro network.

Other analytical model

Research studies conducted by Zhao et al., 2022; Moylan, Foti, and Skabardonis, 2016; Jia, Zhang, and Shi, 2020 focused on data-driven analysis and compared the traffic states before and after the disruption to visualise the impact. Liu, Ma, and Koutsopoulos, 2021 investigated the impact of a metro incident by examining demand changes, passenger accumulation at stations, and travel delays using smartcard and metro AVL data. The impact was demonstrated by a Time Series model. Zhao et al., 2023 using smartcard estimated in-vehicle passenger numbers to capture the impact of bus breakdown on travel behaviour change. The author relied on a Fourier transform method to model the PT travel patterns before and after an incident. Yap et al., 2017 proposed an optimisation model, Transfer Inference Algorithm, with the objective of minimising travel time to capture the mode shift under disruption.

Probability-based approach

The probability-based approach applies stochastic variables that describe the event processes or phenomena that appear to vary randomly. This approach is commonly found in quantitative data analysis where the events or processes, such as the disruption occurrence time or location, are required to be defined by probability distributions. Rusmawati and Rismala, 2016 utilised the Bayesian Network to explore the disruption probability based on train incident log data. Louie, Shalaby, and Habib, 2017 developed an accelerated failure time (AFT) hazard model to describe the relationship between various factors and the resulting delay duration, using incident log data from the Toronto subway system for the year 2013. Saxena, Hossein Rashidi, and Auld, 2019 conducted a survey to explore alternatives during public transport disruptions and generated probabilities for each alternative choice calculated using an Error Component Logit (ECL) model. Zhu et al., 2019 focused on the capacity loss of links due to Hurricane Sandy in New York and created a Monte Carlo model to model the impact. Yap and Cats, 2019 explored a 13-month metro incident log data and identified the probability of each disruption type per station and per time period. In a research study of Dong et al., 2020, the likelihood of node failure was estimated by the Bayes rule, where the failure of a node is assumed to be in relation to the node (intersection) location, flood states, rainfall level and weather.

Wu et al., 2021 described the probability of node failure that was related to a probability of direct failure and cascading failure. Direct failure refers to the actual node failure. The probability of a direct failure was obtained by the historical incident log, while the cascading failure was the failure on sub-nodes passed from the initial directed failed node. The probability of failure for each sub-node is equal, and following this, the static failure cascading can be obtained. Liu, Zhu, and Wang, 2021 introduced a weighted-directed graph to represent regional changes in public transport passenger flow during metro disruptions. For each region (represented by node), the annual number of incidents, the likelihood of incidents occurring on different lines, the occurrence time and the recovery time were calculated based on historical incident data. A Monte Carlo approach and the fitted incident parameter functions were then used to generate stochastically simulated incident events. Rather than fit incident data to certain distribution functions, Mo, Koutsopoulos, and Zhao, 2022 compared travel patterns on normal days versus incident days and proposed a trip cancellation probability to represent the impact of a metro disruption.

Other research begins by assuming a disruption and then uses a probability-based approach to assess its impact. Sun et al., 2016 employed the Bayesian detection method to determine the number of affected PT users and to calculate the probability of passengers choosing to go to a neighbouring station during a disruption based on smartcard data. After processing this data, both pre-disruption and post-disruption scenarios were examined using a simulation model to demonstrate changes in travel reliability. Similarly, Zhao et al., 2021 proposed a demand change ratio based on smartcard data to represent the impact of unplanned events on passenger demand. Cong et al., 2022 employed a logit model to determine the travel behaviour of affected PT users. Yang et al., 2018 estimated the probability of occurrence of a tropical cyclone in Hainan, China, based on the criticality index. Similarly, Yap and Cats, 2022 considered a criticality assessment reflected by passenger betweenness centrality (PBC) to measure the probability of disruption in a multi-modal PT environment.

Regression analysis approach

Owing to causation variables, using the regression analysis approach is able to estimate the relationship between predicted and targeted variables. In transport disruption modelling, this approach is able to describe and predict the severity or impact of disruptions based on the historical disruption data, including the influence factors in relation to the driver, vehicle, road infrastructure and environment (see Zhu et al., 2019; Lovreglio et al., 2020; Yang, Zhang, and Feng, 2022; Tang et al., 2020) or the exposure or frequency of the disruption according to the observed transport states data (see Zhu et al., 2019; Yap and Cats, 2021). Zhao et al., 2021 presented empirical findings demonstrating that event delays are positively correlated with the variability of passenger demand based on extensive smartcard data. Yang et al., 2018 employed a linear regression model to describe the change in bus travel speed due to a subway disruption. Borowski et al., 2023 applied a three-level regression model to define the change in the disrupted ridership over the baseline as the impact of disruption on demand. He, Tao, and Sun, 2024 collected attitudes toward a mode of transport confronting a social event using questionnaire surveys and proposed a regression model to describe the relationship between decision-making on travel mode and social events to show the potential demand change due to a social event.

Machine learning approaches

Due to the abundance of big transport data, data-driven analysis is shifting from conventional statistical models to ML and deep learning algorithms, which can be automatically enhanced based on training data. The statistical models, such as the regression analysis, are preferable in less complex analyses with limited historical data compared to other ML algorithms. Commonly used ML and deep learning algorithms include:

- Random forest (RF) (see Yang, Zhang, and Feng, 2022; Yap and Cats, 2021; Zhao et al., 2021; Tang et al., 2020),
- Dynamic Bayesian network (DBN) (see Wen et al., 2018),
- Back propagation neural network (BPNN) model (see Tang et al., 2020),
- Gradient-boosting decision trees (GBDTs) (see Zhao et al., 2021),
- K-Nearest neighbours (KNN) (see Yap and Cats, 2021; Tang et al., 2020),
- Long short-term memory (LSTM) (see Zhao et al., 2021),
- Multilayer perceptron (MLP) (see Yap and Cats, 2021),
- Support vector machine (SVM) model (see Yang, Zhang, and Feng, 2022; Tang et al., 2020),
- Single-task deep neural network (Single-Task DNN) (see Yang, Zhang, and Feng, 2022),
- Multi-task deep neural network (Multi-Task DNN) (see Yang, Zhang, and Feng, 2022),
- Spatial-temporal Density-Based Spatial Clustering of Applications with Noise (ST-DBSCAN) (see Marra and Corman, 2020),
- Spatial-Temporal Graph Convolutional Network (STGCN) (see Othman et al., 2023).

Proven by previous research studies, the ML-based disruption modelling approach has superior performance to statistical analysis (see Yap and Cats, 2021; Tang et al., 2020). Yap and Cats, 2021 collected PT disruption data containing start time, location and line of disruption occurrence to predict exposure to different disruptions at stations during each time period by two ML algorithms: logistic regression and a Multilayer Perceptron (MLP) classifier. The research study by Yap and Cats, 2021 summarised a framework for analysing the disruption data for predicting the disruptions occurrence and their impacts that have been applied in the majority of related works, where

- predicting disruption exposure by using the supervised ML algorithms, such as logistic regression and MLP, to build the relationship between all available influence factors (such as time of day, day of the week, season, metro lines passing a station, the terminal station, transfer station, passenger volume per station, metro frequency, etc.) and the probability of each disruption type per metro station and per period of time;
- predicting disruption impact by using the supervised ML algorithms, such as the regression, KNN, RF and MLP, to establish the relationship between the available influence factors and the disruption impact subjected to passenger delay per disruption type, per metro station and per period of time;
- establishing a station criticality cluster by using unsupervised ML algorithms, such as hierarchical agglomerative clustering, to develop the relationship between expected annual station criticality per disruption type and period of time and the station clusters.

Zhao et al., 2021 employed the Synthetic Minority Over-sampling Technique (SMOTE) to generate additional observations of disruption events, suitable for machine-learning analysis (RF, GBDT and LSTM algorithms) when predicting the relationship between demand changes and disruption occurrences. Apart from the commonly used ML algorithms, Aparicio, Arsenio, and Henriques, 2021 proposed a discriminative pattern mining algorithm to explore changing urban mobility dynamics before and after the pandemic situation.

Earlier applications of the ML approaches for flood impact were summarised in the review article by Munawar, Hammad, and Waller, 2021.

The major problem when estimating by ML algorithms comes with the data. The quantity and quality of the data limit the quality of the model training and testing, and the quality of data processing or cleansing also similarly impacts the modelling.

2.2.4 Disruption management

Traffic disruptions in transport networks heavily reduce travel reliability and degrade the "image" of the transport system. These negative impacts are unfortunately unavoidable in practice, so the idea of disruption management was born together with the expectation of achieving a more effective, reliable, comfortable and safer travelling environment. The impact management approaches vary widely against the reason for disruptions, which covers rescheduling, rolling stock management, rerouting, signal control methods and bridging transport in order to provide more travelling capacity, as summarised in Table 2.6. By plotting the publications over time, as shown in Figure 2.17, we observe that 2021 to 2022 marks the peak of publishing activity. During this period, more research studies focused on bridging transport, with bus bridging for metro trip recovery being the most frequently mentioned.

Rescheduling

In terms of scheduled PT services, such as buses, trains, trams or metros, the most commonly used service control strategy is rescheduling. This method includes three types of implementation:

- tactical scheduling, which is the published planned timetable that is followed by the operators;

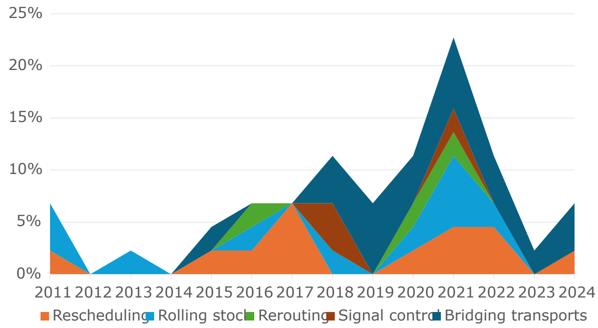


FIGURE 2.17: Time series of various disruption management approaches proposed in the literature.

- operational scheduling, which is used for temporary or auxiliary services for congestion mitigating;
- rescheduling, which is used when emergencies or disruptions occur, and further remedy is expected.

The objective function required by the rescheduling approach often aims to minimise the travel delay and time needed to schedule or dispatch. The existing body of disruption management for railway networks has provided many comprehensive studies in relation to the rescheduling approach. Recent evidence provided by Veelenturf et al., 2016 applied the integer linear programming model to reschedule the train services that were blocked on one or all tracks within the study area for up to two hours. This research took large-scale disruption (full track block) into account and verified the feasibility of the method by conducting scenario-based experiments. Following the concept of cost minimising, Yang et al., 2019 developed a new approach that also took the energy cost and delay disturbances into account when rescheduling metro services.

Rolling stock

The other major management approach that is commonly seen in rail networks is rolling stock. Since the crews are coming with the vehicles, the rolling stock method is sometimes combined with the crew scheduling. Therefore, the cost of work, as well as the cost of labour on the rolling stock approach, tend to be less attractive than the simple, practical and economic rescheduling approach. A comprehensive research study on rolling stock has been undertaken by Wang et al., 2021b with a full review on the subject of rolling stock planning, scheduling and also application.

Rerouting

The methods of rerouting stem from the trip assignment approach, where the ideal trip assignment should maintain minimum travel cost, which requires all assigned trips to have the minimised travel cost. Compared with the trip assignment in a normal situation, the impact on road capacity should be considered in the scenario when traffic disruption occurs. The idea of rerouting should be subjected to the total travel time and the maintenance of travel purposes, as in practice, passengers need to reach their destination as fast as possible. Dobler and Nagel, 2016 proposed a method of within-day re-planning for simulating trip reassignment processes based on the information of traffic disruption.

TABLE 2.6: Various disruption management approaches involved in the literature

Management approaches	References
Rescheduling	Yap et al., 2017; Yap et al., 2022; Wang et al., 2021b; Leng, Liao, and Corman, 2020; Zhang et al., 2024b; Placido, 2017; Corman, 2020; Meng and Zhou, 2011; Zhan et al., 2015; Veelenturf et al., 2016; Ghaemi, Cats, and Goverde, 2017; Li et al., 2021
Rolling stock	He et al., 2021; Zhang et al., 2022; Wang et al., 2021b; Leng, Liao, and Corman, 2020; Dekker and Panja, 2021; Veelenturf et al., 2016; Nielsen, 2011; Melo et al., 2011; Cadarso, Marín, and Maróti, 2013; Schipper and Gerrits, 2018
Rerouting	Leng, Liao, and Corman, 2020; Veelenturf et al., 2016; Luo and Xu, 2021
Signal control	Qi, Zhou, and Luan, 2018; Mihăiţă, Dupont, and Camargo, 2018; Mao et al., 2021
Bridging transports	Diab, Feng, and Shalaby, 2018; Tan et al., 2020; Zhang et al., 2024a; Liu, Ma, and Koutsopoulos, 2021; Liang et al., 2019; Kepaptsoglou and Karlaftis, 2009; Jin, Teo, and Odoni, 2016; Xu and Liang, 2020; Tang et al., 2021b; Shao and Song, 2022; Tan et al., 2023; Chen et al., 2024; Yin et al., 2018; Wang, Yuan, and Yin, 2019; Wang, Jin, and Sun, 2022; Luo and Xu, 2021; He et al., 2019

The transport system was modelled by an agent-based model in MATSim, and the trip assignment was calculated once to simulate the reaction of passengers against a current incident event. A similar application of this rerouting control has been found in the research study by Rahimi and Corman, 2020, who proves its effectiveness on journey recovery and reduction of total delay by re-planning the trips in a multi-modal PT network. **<empty citation>** considered both the pre-disruption and post-disruption processes and proposed a rerouting model based on iterative trip reassignment optimisation against the impacted trips, non-impacted trips and time of clearance in rerouting.

Signal control

Signal control plays a vital role in balancing the flow and capacity of the intersection by modifying the timing plans. This method is often required to be used as the real-time approach that is able to adjust the timing by real-time traffic states or temporarily nearby incidents. Similar to the rescheduling approach, the signal control also belongs to the stochastic optimisation-based approach that is subjected to minimising the variables such as travel cost, travel time, delay or length of the queue. The signal control has limited ability to adjust traffic density because of the limited buffer time available to be used in the timing plan, while an overlong buffer time would decrease the intersection capacity. Therefore, such a control approach is often used for preventing congestion and will assist in releasing crowding. With the development of telecommunications for storing, retrieving and sending information, real-time signal control has been encompassed by transport control.

Bridging transport

Bridging transport is often used as a response plan after long-lasting railway disruptions when there is a need to reconnect the origin and destination train stations with an alternative, which normally refers to the buses. In common practice, bus bridging is a pre-defined response strategy considering the disruption location, duration, number of impacted trips, available bridging buses or shuttles, available labourers and road conditions when designing the bus bridging plan that maintains the minimum travelling cost for both operators and passengers. The general framework of a bus bridging plan includes:

- Step 0. Initialise the network and identify the impacted paths and train stations.
- Step 1. Generate the candidate set of bus bridging routes and service frequencies.
- Step 2. Calculate the expected total travel cost and unsatisfied travel demand of the candidate routes.
- Step 3. Identify the best bus bridging route and alternative bus bridging routes.
- Step 4. Evaluate the best route following the four-step travel model.
- Step 5. Update the total travel cost and unsatisfied travel demand as the control performance.
- Step 6. If the control performance satisfies the pre-defined criteria, then stop; otherwise, select the best route from the set of alternatives from Step 3, then go to Step 4.

Among publications on disruption response plans, many focus on bus bridging. Some research studies have worked on scheduling feeder buses using optimisation procedures, such as the column generation solution (see Liang et al., 2019; Jin, Teo, and Odoni, 2016; Xu and Liang, 2020; Shao and Song, 2022; He et al., 2019). Alternative travelling routes for bus-disrupted commuters are generated using a multi-modal k-shortest path model (see Luo and Xu, 2021), and heuristic algorithms have been used to investigate response plans (see Tan et al., 2020). Earlier contributions regarding the bus bridging approach were summarised in a review article by Zhang et al., 2020.

2.2.5 Discussion and future directions

Findings on disruption causes and impacts measurement

In subsection 2.2.2, an overview of the disruption causes and the methods of measuring their impact have been presented. Based on the related publications, the causes of traffic disruptions can be divided into two major categories, namely natural disasters and human-related failures. Both of them can result in the disruption of travelling by obstructing or damaging the access. For most of the disruptions due to natural disasters, the costs of the impact are often more than that of the man-made incident. The cost of travel disruption typically includes the impacts on travel efficiency, society or the economy.

Disasters occur unpredictably, resulting in limited cases for study. Numerous publications have shown that floods and storms are more frequent worldwide, leading to more research on these topics. Such disasters primarily affect road function and infrastructure, with studies often focusing on the impact on link or node capacity from a network perspective. Reviews indicate that research on cars has garnered more attention than public transport, with more causes related to natural disasters and transport disruptions, such as fires or sea level rise, being discussed. However, there is no literature addressing the impact of fires or sea level rise on public transport. Apart from the causes, the impact on car function was also discussed. For example, research conducted by Pregnolato et al., 2017 explored the travel speed reduction in relation to the depth of flood. However, there is a lack of research on the impact of such factors on PT travel speed.

Some works have measured the disruption impacts on road networks by using data-driven approaches, but those research studies tended to assume that the impacts followed a uniform distribution or the impacts were maintained static. Some research has studied the impact of car travel speed reduction and identified a bathtub model, which divides the impact into three stages: the starting phase, requiring a transition plan; the disruption phase, where travel is impacted; and the end phase, where a recovery plan helps restore the network to normal. Such analysis methods have been applied to railway networks, as discussed in the review by Ghaemi, Cats, and Goverde, 2017. This detailed disruption modelling approach has the potential to be applied to other networks as well.

The previous research studies have captured mostly the impact propagation from the space point of view; few of them modelled the change of impacts in time. There is little published data on large-scale disruptions due to natural disaster events, and the relevant quantitative analysis of the disruptions is mostly based on traffic state data, such as travel speed and traffic flow, for a single network. There is a lack of analysis and comparison of the impact of a single disruption on a single network, a single disruption on multiple networks, multiple disruptions of different networks on multiple networks, etc. There is still a gap when it comes to the clear categorisation of the impact of disruptions on PT networks with regard to time and space.

Furthermore, existing research has not extensively explored the impact of disruption on different modes of PT. This area presents a potential direction for future research, particularly regarding transfer inference. Some research has considered the idea of transferring from rail to road transport during disruptions, mainly focusing on metro-to-bus transfers. More solutions could be included, such as shared bicycles, e-bicycles, e-scooters, cars, or other on-demand private or public transport options.

Over time, research on the impact of disruptions on PT has shifted from changes in the network reliability or supply to passenger demand and travel patterns. Particularly due to the pandemic, concerns about health impacts have influenced mode or route choices, highlighting the need to explore short-term adjustments in transport supply, and the power of lead by propaganda in order to leverage the PT infrastructure wisely.

Apart from disruption impact identification, according to previous studies, the definition of small and large-scale disruption is still vague. Some research studies referred to the closures of multiple roads as a large-scale disruption. Others define large-scale disruption as a disruption that lasts for a long time and impacts traffic for a long time. These nebulous definitions raise a concern when it comes to qualifying the disruptions by their temporal and/or spatial impacts. Much of the research up till now has not been able to draw on any systematic research into establishing the comprehensive causation or correlation between various causes and their consequence under large-scale disruptions. A large-scale disruption can be one that impacts a large area, multiple links or nodes; it can also be one that lasts for many hours; it can also be one that a large number of passengers, a large number of vehicles, properties or other fixed objects is affected by; the consequences of the impact on society and economy can be used as the criteria; the recovery time and cost or the recoverable damage to the transport network can also be considered as the standards. The work of defining disruptions assists in the categorisation of the events and saves effort when selecting a proper response plan accordingly.

Findings on disruption modelling methodology

So far, network modelling and disruption modelling have more capability in imitating real-world travelling and transport processes. Assisted by advanced information and computing technologies, the task of modelling has been led by simulation and data-driven methods. Multiple computer-aid software tools have also become accessible to mitigate the workload on data processing and modelling. The development of the network modelling approach failed to encourage disruption modelling as most of the disruption modelling was limited by a node or link removal based on the graph theory. There were empirical evidence-based initiatives on treating disruption modelling as a process consisting of three phases: pre-disruption, disruption, and post-disruption. The impact before and after the occurrence of the disruption also matters. Moreover, the propagation of the disruption has only been achieved based on the concept of failure cascading, and there is still a gap in propagation estimation with the consideration of traffic and movement states in the network.

Both the objectives and methods used for network modelling depend on the type and quality of the available data. Due to limitations in data quality and quantity, most disruption modelling in the literature focuses on a single transport mode network. Furthermore, constraints in modelling and programming methods often prevent multi-modal network-based disruption modelling from effectively combining the impacts on road and public transport networks. This is particularly true for networks isolated to specific modes, such as train or metro systems. However, for bus networks that share links with general traffic, the impact between cars and buses can be more easily considered.

Some studies, such as those reviewed by Pregnolato et al., 2017; Jarmuz and Chmiel, 2020), use time-dependent travel speed to identify impacts, a method often applied to car disruptions. Conversely, Zhao et al., 2023 found that disruptions reflected in smartcard data are less apparent. This research focuses on modelling public transport patronage patterns by analysing changes in swipe data to identify reductions in counts due to disruptions. In modelling disruption impacts, data quality and quantity are crucial. However, if a public transport network is designed to resist disruption propagation, how does this affect the analysis? For research studies focusing on network vulnerability or resilience, a promising future direction could be to evaluate how robust the network remains even during disruptions.

Current literature shows that network modelling has been addressed according to a variety of concepts or theories, such as the concept of cost matrix, graph theory, probability theory, concept of optimisation, simulation or the concept of data-driven analysis. The modelling approaches derived from the above ideas have a growing ability to illustrate reality as increasing numbers and levels of detail have been able to be considered in a model. However, this development in network modelling disregarded the modelling of network disruptions. The vast majority of the work on disruption modelling has been limited to instant nodes or link removal when modelling rather than following the fact of temporally and spatially gradual changes in impacts. A few studies simply considered the scenario-based approach to uncover the change in disruption impact in time or space. These scenario-based experiments provided a good foundation when it comes to understanding the impact of change in time and space, but the reliability of the result highly depends on the quality of modelling and how the network is built.

Assuming that the disruption results in road closures for a period of time, the trip reassignment to the network that follows the disruption is the most commonly used method when examining the temporal and spatial difference in disruption impact on network performance. Such a scenario-based approach also has the potential to illustrate the impact propagation; for instance, the hypotheses on the location of road closures can be arranged radially, either following the connected links or diverging in lines from a common centre. The closure of the road can also be replaced by a road degrading. The degrading level is also manipulable and is also available when representing the impact propagation in the network.

Findings on disruption management

Analysis by scenario-based approaches includes manipulating a variable to determine the exact cause that changes another variable or leads to a specific consequence. This method heavily depends on the control variable and the relevant hypothesised scenario defined in the experiment. By comparing the result of each scenario to that of the base scenario, it is possible to uncover the relationship between cause and result. The base case scenario often refers to a set of basic assumptions that are supposed to bring about the most realistic representation of an event. In traffic disruption modelling, the scenario-based approach often considers an intervention plan that mitigates the negative impact of transport disruptions on the networks. The most common control approaches that appear in the literature consist of rescheduling, rolling stock, rerouting, signal control and bridging transport. These are operation-level approaches, which balance the level of supply to release the demand in a disrupted network.

The previous studies mentioned the need to clearly and systematically identify the impacts of disruption, which leads to another gap in guiding the match of the proper intervention plans towards each type or scale of disruption. Transport infrastructure is limited in its capacity; thus, a plan ahead that leverages the existing resources when maximising the performance of the transport network and journey recovery becomes way more important.

Current intervention plans include the re-scheduling of PT services to mitigate the impact of disruption. This method offsets the disruption cost from the operation side, which also involves the arrangement of the rolling stock and labour re-scheduling. A method considering the point of view of the passengers is re-routing for the trips. This method comes from the natural behaviour of a passenger to save the travel cost when confronting a journey disruption. The other common method that seems far from the transport users and operators is signal control, where the buffer time is employed to cope with the unbalanced traffic flow in both directions. Bridging transport is often used when large transport disruption occurs, but there is a great demand between the origins and destinations. However, the extra vehicles and labour costs and the costs of route redesigning always force operators to think twice before implementation.

A number of previous studies demonstrated the effectiveness of using bridging transport for journey recovery. However, there remains a need to clarify the potential increment in the delay on roads due to the extra load added by the bridging transports. In a multi-modal PT environment, passengers are provided with more travelling methods. Passengers can choose to wait for the service recovery, transfer to another transport mode, or expect bridging transport. Therefore, when the operator should consider using the bridging transports and how to route them in multi-modal PT networks becomes a next-level challenge.

2.3 Conclusions

The main motivation of transport modelling is the requirement of transport planning, management and operation. There is a need for a testbed to uncover the current network deployment, infrastructure usage and travel pattern. The transport network modelling enables the network deployment to be illustrated in detail to display the topological features as well as the geometry parameters of the network.

As mentioned in this review chapter, three key network modelling methods were mostly mentioned in the literature, namely the Graph theory-based modelling approach, layer-based modelling approach, and Supernetwork approach. The layer-based modelling approach has originated from the Graph theory approach, where the Graph theory approach produces one graph that represents a network, all network elements are included in this single graph, while the layer-based approach generates multiple graphs that represent different elements of a network and each graph is called a layer of the network. In some cases, each layer contains the integrated configurations of a network, in this way, a layer-based modelling approach can demonstrate the multi-modal transport networks or combination of a network model and other models, such as the demand model.

According to the review, the powerful modelling approach as the layer-based approach has been limited to represent a single public transport network in most of the literature. The research studies have indicated the success in displaying the network interplay, such as transferring (e.g. walking to public transport stations), by using such a layer-based approach. Such the attemption on uncovering the interplay can be further broadened by adding more transport mode' networks so that the interaction between different transport mode' networks can also be investigated. Thus, the demand propagation within a multi-modal public transport network can be estimated at a macroscopic level. More efforts can therefore focus on the modelling of the interplay between different transport modes' networks. For instance, in a traditional Space P network representation, distinguishing the weight of edges that connect two nodes in the same network and the weight of edges that connect nodes in different networks can be a new research question. When considering the layer-based modelling approach, the layer for representing the transferring between nodes that are in the same network need to be separated from the representation of transferring between different networks.

The Supernetwork modelling approach is essentially an agent-based modelling approach. The Supernetwork approach focuses on the individual travelling; it has a strong ability in representing the trip chain. However, individual trip modelling at the microscopic level cost much in time, but the concept is worth keeping when mesoscale or macroscale modelling. Future research can focus on the quantification of transferring between different networks or the interaction of different agents within a network.

With the availability of the real transport data, researchers can calibrate the network model better; the definition of the edge weight in Space L, for example, might largely benefit from the real data. So does the weight for transferring or other information in relation to the interplay within the same network or between different networks.

The quality of the demand modelling could also benefit more from the real data. Most of the method was proposed to estimate the demand due to the lack of such data. However, the story might differ when infrastructure usage data are unrestricted: real data can be used to calibrate the OD estimation model, it can also be used to train the model and discover the new and more accurate relationship between the OD matrix and its influence factors. Among the publications, the forms of the deterrence function were proposed based on experiences, and there is still a gap in testing and comparing the different forms of the deterrence function by using the real data.

Regarding the demand modelling methodology, the travel behaviour between different networks in the mesoscopic and macroscopic levels are still underestimated, compared with the microscale modelling. The mode choice and the decision-making when transferring in a multi-modal public transport network are maintaining lacking in exploring. Understanding the nature of the decision-making is only halfway towards success; the method of accurate behaviour prediction is the solution that needs to be chased.

In recent decades, the structure of public transport has changed from a single network and separate networks to integrated multiple networks. Along with the change in advancing mobility, it is also more likely to witness many disruptions in such multi-modal public transport networks due to their complexity. Since networks are connected, the disruptions can potentially propagate through joined nodes or links in a multi-modal public transport network, and in effect generate large-scale disruptions that require more effort to address. In the literature, large-scale incidents, such as natural disasters or man-made catastrophes, are the major causes of large-scale disruptions. The impact of disruption propagating in a complex transport network, as well as the resulting post-disruption, has become an increasingly important problem. This chapter presents a review of the published literature discussing multi-modal public transport network modelling and disruption modelling. The existing contribution towards the subjects is categorised by particular perspectives, namely, modelling approach, disruption reason and impact on the network, and relative intervention plans.

The existing literature demonstrates that network modelling has been approached through various concepts and theories, such as the cost matrix concept, graph theory, probability theory, optimization concept, simulation, and data-driven analysis. These modelling approaches have improved their ability to depict reality by considering an increasing number of details and levels of complexity in the models. However, there has been a lack of emphasis on modelling network disruptions. The majority of research in disruption modelling has focused on the instant removal of nodes or links rather than capturing the gradual temporal and spatial changes in impacts. Some studies have employed scenario-based approaches to explore the evolving impact of disruptions over time and space. While these experiments provide valuable insights into the temporal and spatial changes, the reliability of the results depends heavily on the quality of the modelling and network construction. The most commonly used method to analyze the temporal and spatial differences in disruption impact on network performance is trip reassignment to an alternative network following the disruption. This scenario-based approach also allows for illustrating impact propagation, such as arranging road closures in radial patterns, following connected links, or diverging from a central point. Road closures can also be simulated by degrading road conditions, with the level of degradation being adjustable and facilitating the representation of impact propagation in the network.

Previous research has primarily focused on railway networks, with most disruptions involving link closures. However, road networks are more susceptible to link degradations in practical scenarios. The situation is similar for multi-modal public transport networks, where different transport modes can collaborate to handle increased demand, allowing for dynamic changes in road capacity. Nevertheless, there is still a gap in measuring the impact of disruptions on multi-modal public transport networks, as previous studies mainly concentrated on single transport networks without considering impacts on other modes of transportation. Data-driven approaches offer a more reliable and accurate analysis. By utilizing real-world data that captures traffic flow characteristics such as density, flow, speed, and volume, researchers can establish relationships between variables such as time or location and the consequences of disruptions. This data can aid in determining the extent of link degradation by assessing changes in traffic flow. Furthermore, the availability of real-world data enables the calibration of transport demand models, facilitating the assessment of disruption impacts on other transport modes networks. One possible approach is to employ the classic four-step model, considering network configurations as variables. By assigning trips to different networks and observing the decrease in traffic flow or passenger travel time delays, one can infer changes in network impact. Another valuable application is the scenario-based approach.

Apart from this, according to previous studies, the definition of small and large-scale disruption still needs to be clarified. Some research refers to the closure of multiple roads as a large-scale disruption, while others define it as a prolonged disruption with long-lasting impacts on traffic. These vague definitions create challenges when categorizing disruptions based on their temporal and spatial effects. Previous research has yet to systematically establish a comprehensive understanding of the causation or correlation between different causes and their consequences during large-scale disruptions. A large-scale disruption can be characterized by its impact on a wide geographical area, multiple links or nodes, long duration, significant effects on passengers, vehicles, properties, or fixed objects, as well as its societal and economic consequences. Recovery time, cost, and the extent of damage to the transport network can also be considered criteria for defining large-scale disruptions. As mentioned in the introduction, large-scale disruptions often stem from major incidents such as natural or man-made disasters. Additionally, the direct reason for the disruption can be used as a criterion, regardless of its socioeconomic impact. Establishing clear definitions for disruptions aids in event categorization and facilitates the selection of appropriate response plans.

The previous studies mentioned that the need to clearly and systematically identify the impacts of disruption, which leads to another gap in guiding the match of the proper intervention plans towards each type or scale of disruption. Transport infrastructure is limited in its capacity; thus, a plan ahead that leverages the existing resources when maximising the performance of the transport network and journey recovery becomes way more important. Current intervention plans include the re-scheduling of public transport services to mitigate the disruption impacts. This method offsets the disruption cost from the operation side, which also involves the arrangement of the rolling stock and labour rescheduling. A method considering the point of view of the passengers is re-routing for the trips. This method comes from the natural behaviour of a passenger for saving the travel cost when confronting a journey disruption. The other common method that seems far from the transport users and operators is signal control, where the buffer time is employed to cope with the unbalance of traffic flow in both directions. Bridging transport is often used when large transport disruption occurs, but there is a great demand between the origins and destinations. However, the extra vehicles and labour costs and the costs of route redesigning always force operators to think twice before implementation. A number of previous studies demonstrated the effectiveness of using bridging transport for journey recovery. However, there remains a need to clarify the potential increment in the delay on roads due to the extra load added by the bridging transports. In a multi-modal public transport environment, passengers are provided with more travelling methods. Passengers can choose to wait for the service recovery, transfer to another transport mode, or expect bridging transport. Therefore, when the operator should consider using the bridging transports and how to route them in multi-modal public transport networks becomes a next-level challenge.

Although previous research studies have already covered the traditional topics related to general transport modelling, there are still many open questions about a standardised methodology to deal with multi-mode transport disruptions. The modelling methods need to be improved, the mechanism of disruption still needs to be uncovered via real data, the disruption details need to be clarified, and the response plans need to be categorised. Future studies may consider these requirements and extend the topic to a more complex multi-modal transport network modelling.

Chapter 3

Origin-destination matrix estimation for public transport

Estimating the large-scale Origin-Destination (OD) matrices for multi-modal public transport (PT) in different cities can vary largely based on the network itself, what modes exist, and what traffic data is available. In this chapter, to overcome the issue of traffic data unavailability and effectively estimate the demand matrix, we employ several data sets like the total boarding and alighting, smart card as well as the General Transit Feed Specification (GTFS) in order to capture the PT dynamic patronage patterns. First, we propose a new method to model the dynamic large-scale stop-by-stop OD matrix for PT networks by developing a new enhancement of the Gravity Model via graph theory and Shannon's entropy. Second, we introduce a method entitled "Entropy-weighted Ensemble Cost Features" that incorporates diverse sources of costs extracted from traffic states and the topological information in the network, scaled appropriately. Last, we compare the efficiency of a single travel cost versus various combinations of travel costs when using traditional methods like the Traverse Searching and Hyman's method, alongside our proposed "Entropy-weighting" method; we prove that our method coupled with a multi-modal PT OD matrix modelling is superior to traditional methods.

This chapter is based on an edited edition of the following article: Zhao D, Mihaita AS, Ou Y, Grzybowska H, Li M. Multi-model Origin-Destination Matrix Estimation for Public Transport via a Weighted Graph and Traffic Features Integration. Under review at the Journal of Transportation research part C: emerging technologies.

Contents

3.1 Introduction

3.1.1 Background and motivation

Public Transport Origin-Destination matrices (PT OD), also known as trip matrices, serve as a foundational element for traffic demand estimation, providing insights into how the travel demand is distributed in space and time. This information is crucial for planning, designing and managing transportation systems in a way that is responsive to the actual needs and behaviours of the population.

Therefore, several techniques for trip matrix estimation have been attracting scientific attention since the last century (Balcombe et al., 2004).

The choice of what OD matrix estimation model should be adopted largely depends on the data availability. The smart card data (Hussain, Bhaskar, and Chung, 2021), mobile phone tracking (Ge and Fukuda, 2016; Wismans et al., 2018) or Bluetooth and Wi-Fi tracking data (Ta et al., 2018; Huang, De Villafranca, and Sipetas, 2022) provide comprehensive information on each trip from an origin stop to a destination stop following a certain PT route, which enables it to be one of the ideal data resources for PT demand modelling. However, such extensive data may not be accessible in many cities. In cases where full trip matrix data are unavailable, the total trip generation and attraction (GA) data are often derived from census data or surveys (Tamblay et al., 2016). Alternatively, aggregated data from smart card transactions is often utilised, as demonstrated in this chapter for practical considerations.

Given the total trip GA from smart card transactions, there is a need for an estimation model to derive the OD matrix. The Gravity Model is a conventional modelling approach suitable for this purpose (Ortuzar S. and Willumsen, 2011). Originating from the gravity law, this model was first introduced in 1931 by William to delineate the connection between trip distribution and microscopic zonal demand (GA). The accuracy of the estimated OD matrix hinges on the model's calibration; therefore, subsequent to the introduction of the Gravity Model, numerous endeavours have been made to enhance the calibration of this model.

Past studies on the Gravity Model have two main categories for model calibration: a) the first involves the calibration of balancing factors, which are weights against the total generation and attraction in the model. These factors aim to ensure that the model accurately reflects the real-world travel volume; b) the second category focuses on calibrating the deterrence function, which is a model that represents the resistance or travel cost associated with travelling between OD pairs. By calibrating this function, transportation planners can consider various factors influencing travel behaviour, such as travel time, distance, and speed. However, due to limitations in the travel cost data, different calibration methods have been proposed to compensate for biases in the travel cost features, reflect travel preferences, and enhance the model performance. This chapter delves into the relationship between the form of the deterrence function and the travel cost features, aiming to deepen our understanding of this crucial aspect of the model calibration.

In the existing literature, travel cost features primarily focus on traffic-related factors, see Literature Review in Chapter 2. Only one research study, conducted by Rubio-Herrero and Muñuzuri, incorporated fuel, driver, tyres, and maintenance costs into OD estimation for freight networks. However, this study did not assign weights to each of these travel cost features. The inclusion of multiple travel cost features presents another calibration challenge: determining the weight assigned to each feature. To address this gap, we introduced an entropy-ranking method to determine the weight for each travel cost feature effectively. This enables the model to derive maximum benefit from multiple travel cost features, thereby enhancing overall accuracy.

In our work presented in this chapter, we also consider, in addition to the traffic states, that the trip distribution is also associated with the network topology because the geometric properties can affect the node and the link capacity, as well as the passenger accessibility in the network; this can further influence the link travel capacity and the travel efficiency. Especially when considering a dynamic cost matrix, the pre-defined timetable can interfere with timely access for travellers in specific locations.

Therefore, this intuition raises our initiative to employ the topological features for the cost matrix estimation as well, as this will reflect the network accessibility and the inter-modality between all modes inside an interconnected PT graph. The graph theory-oriented features have not been explored extensively in the past, the exceptions being some recent studies regarding network vulnerability analysis; only one recent study has explored their potential by separately using the betweenness and the travel time as travel costs (see Lu, 2018). Furthermore, by using the Gravity Model, we conduct a comparative analysis to assess the efficacy of various travel cost features in OD estimation. This evaluation encompasses the utilisation of both traditional Traverse Searching and the Hyman method, as well as our proposed entropy-ranking method for calibrating the deterrence function.

In line with the method of weighting by the friction factor, the entropy measurement has the potential to estimate accurately the friction factor (see Ai, 2017). Entropy was initially proposed in the information theory as a measure of the uncertainty or randomness in a system, it can be used in order to quantify the degree of variability or diversity for each travel cost feature. It measures the average amount of information required to draw an outcome from a probability distribution. Therefore, by ranking the travel cost features based on their entropy, it is able to identify the factors that are most important in shaping travel behaviour. Once we have ranked the travel cost features by entropy, we can integrate various impacts into one index, which can be used to calibrate travel patterns in a Gravity Model. This approach allows us to capture the overall impact of different travel cost factors on the travelling pattern and to develop more accurate and reliable models for estimating travel demand. So far, the entropy measure has only been used in ranking or for evaluation purposes in the current literature (see Zhang and Ng, 2021; Ai, 2017; Qi et al., 2021; Wei et al., 2022); to the best of our knowledge, no publication indicates the feasibility of using the entropy to rank the travel cost features and the advantage of using the entropy weighted cost matrix in the OD estimation.

To fill the gap in PT OD estimation, this chapter introduces a framework for dynamically estimating the stop-by-stop OD matrix for large-scale PT networks. We enhance the traditional Gravity Model by incorporating entropy-weighting on travel cost features. Our study investigates the influence of various PT network travel cost features on OD estimation, encompassing traffic characteristics such as travel time, travel distance, and fare cost, as well as graph topological features like connections, closeness and straightness. We evaluate the performance of different forms of deterrence function used in the Gravity Model, and we compare traditional calibration methods like Furness and Hyman against our proposed Entropy-weighted method. The results demonstrate significant improvements in terms of RMSE, MAPE and MAE.

In prior research studies, the deterrence function was typically associated with a single type of travel cost due to the challenges of determining the optimal weightings for multiple cost features (Tamblay et al., 2016; Suprayitno, 2018). Some studies incorporated multiple travel costs but without assigning weights to each cost feature (Rubio-Herrero and Muñuzuri, 2021). Expanding upon previous research, in this research chapter we introduce a new innovative approach to rank and assign weights to various travel cost features more efficiently. By incorporating multiple travel cost features into the Gravity Model, we aim to significantly enhance the model's overall performance and show its efficiency as compared to more traditional techniques.

3.1.2 Challenges and contributions summary

In order to obtain accurate public-transport-related OD matrices that are multi-modal, we encounter several challenges (open questions):

- How to dynamically model the spatial and temporal travel patterns of PT passengers?
- How to effectively obtain accurate OD matrices for multi-modal PT networks using a minimum data availability?
- What factors have a critical influence on the PT OD estimation across various types or modes of the PT networks?

To address these challenges, in our current work we propose:

- a new framework to dynamically estimate the stop-by-stop PT OD matrix for multiple transport modes;
- an extended version of the Gravity Model for our multi-modal PT OD estimation. This relies on a novel model calibration method using Entropy-weighting, which considers both the traffic characteristics (travel time, travel distance, and fare cost) as well as the graph topological features (connections, closeness and straightness);
- and finally, we show that our new proposed PT OD calibration method outperforms the traditional approach, namely Hayman's and Traverse Searching methods, in terms of RMSE, MAPE and MAE.

3.1.3 Organisation of chapter

The rest of this chapter is organised as follows: in Section 3.2 we first present the framework of the large-scale stop-by-stop OD estimation for PT. Following the framework, the details of the extended Gravity Model are further highlighted in Section 3.2.2, where we show the modelling method for aggregated GA vectors (see Section 3.2.2) and deterrence function considering "single travel cost feature" and "multiple travel cost features". Then, the classical deterrence function calibration processes are discussed in Section 3.2.2 against the proposed method of "Entropy-weighted" in Section 3.2.2. Finally, the topological features' modelling methods are included in Section 3.2.2. The performance of an OD estimation is validated using MAE, RMSE and MAPE, shown in Section 3.2.3. The application of the estimation methods involving the three deterrence function calibration methods to a real network is presented in Section 3.3 with results shown in Section 3.4, where The ideal parameter configurations for the deterrence function and the most appropriate form of the function tailored to various travel cost features are presented in this. Finally, Section 3.5 is provided to clarify the research limitations and offer future directions in this field.

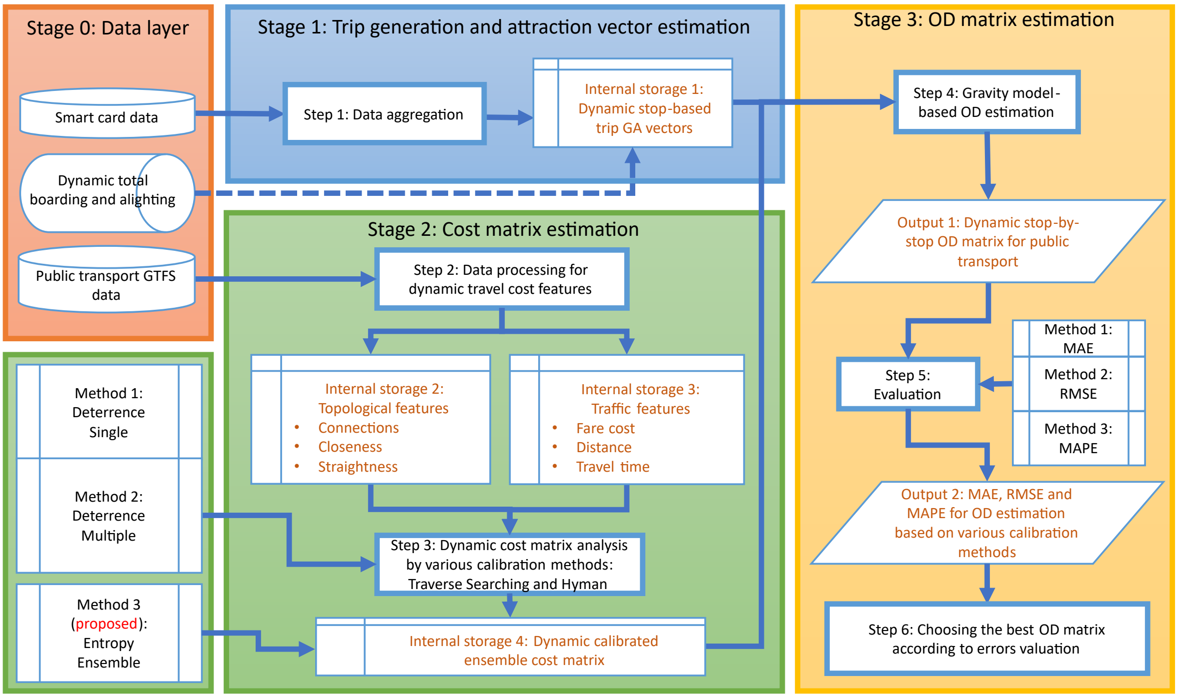


FIGURE 3.1: The framework of our proposed dynamic stop-by-stop OD estimation for large-scale multi-modal PT.

3.2 Methodology

3.2.1 Modelling framework

The framework presented in Fig. 3.1 illustrates our approach to dynamically estimating the stop-by-stop OD matrix. This framework comprises three stages. In *Stage 0*, we collect, filter, and clean all input datasets, including smart card data and Global Transit Feed Systems (GTFS) data. At this stage, we also integrated total counts of boarding and alighting. This integrated data serves as an ideal alternative for GA vector generation in *Stage 1*; *Stage 2* illustrates the deterrence function calibration process considering three types of data resources: the "Deterrence Function Considering Single Cost Features", "Deterrence Function Considering Multiple Cost Features" and "Entropy-weighted Ensemble Cost Features". When it comes to individual travel cost features, we employ two calibration methods, namely Traverse Searching and Hyman's methods. Conversely, when dealing with multiple costs, we adopt a fusion approach by utilising the Hyman method alongside our proposed Entropy-weighted method to demonstrate the effectiveness of OD estimation. Our proposed methodology leverages entropy ranking to assess and weigh various travel cost features, aiming to optimise the accuracy of the deterrence function and, consequently, OD estimation. For each calibration process, we evaluate the efficacy of three function forms: power, exponential, and Tanner; at the *Stage 3*, we employ the Gravity Model for the OD estimation and further validate the feasibility of the proposed deterrence function calibration method. At this final stage, the best cost matrix estimation method and the most accurate OD matrix are selected according to MAE, RMSE and MAPE.

3.2.2 PT OD Estimation using the Gravity Model

The Gravity Model

For PT such as buses or trains, the unit of its OD matrix is defined as the number of passengers, and the content of the OD matrix is the total number of passenger trips. Unlike the OD matrix for cars, the origins and destinations for the PT network are the PT stops. *Adjacency Matrix:* The existence of trips

between each OD pair depends on the predefined paths for routing the PT in the network. Therefore, the PT network is defined as a Space L' graph, where the nodes are represented by the PT stops and the links between nodes are the routes between any PT stops. In a graph of Space L', different edges represent different links used in different networks with different directions Yang et al., 2014. To capture the nature of dynamic timetables, such information is recorded in a time-dependent adjacency matrix:

$$H(t) = [h_{m_i,n_j}(t)], i, j = \{1 \dots J\}, m, n = \{1 \dots N\}, \tag{3.1}$$

where J is the total number of statistical areas inside the sub-network; N represents the total number of PT stops in the network; and

$$h_{i,j}(t) = \begin{cases} 1, & \text{if } (i, j) \text{ is an accessible link at t,} \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

The Gravity Model: The number of time-dependent trips made by a public transit mode *pt* is obtained

based on the placement and existing routing between PT stops, where *pt* is considered as buses (b), train and metro (tn). Therefore, the total number of trips departing from an origin stop $m \in \{1 \dots N\}$ to a destination stop $n \in \{1 \dots N\}$ can be calculated based on the Gravity Model with a given total number of trips departing from the origin stop m by a mode *pt*, denoted by $o^{pt}_{m_i}$, and that of trips arriving at a destination stop n by mode *pt*, denoted by $d^{pt}_{n_j}$. The PT trip matrix can be represented by $OD_{pt}(t) = [od^{pt}_{m_i,n_j}(t)], i, j = \{1 \dots J\}, m, n = \{1 \dots N\}$. The total number of trips departing is also known as the trip generation. In contrast, the total number of trips arriving is known as the trip attraction in the traditional four-step OD estimation. The most common form of the Gravity Model is expressed as:

$$od^{pt}_{m_i,n_j}(t) = A^{pt}_{m_i}(t)\ o^{pt}_{m_i}(t) B^{pt}_{n_j}(t) d^{pt}_{n_j}(t) f\left(c^{pt}_{m_i,n_j}(t)\right), \tag{3.3}$$

where $A^{pt}_{m_i}$ and $B^{pt}_{n_j}$ represent the time-dependent weights towards the total number of origins ($o^{pt}_{m_i}$) by PT and the total destinations ($d^{pt}_{n_j}$), respectively; $f\left(c^{pt}_{m_i,n_j}\right)$ is the travel cost function that represents the time-dependent travel cost between two zones by a mode *pt*. three cost matrix calibration methods are compared in this chapter, as detailed in Section 3.2.2, Section 3.2.2 and Section 3.2.2. The origin stop m belongs to the origin zone i while the destination stop n belongs to the arrival zone j. *The*

constraints: Two constraints are employed in the Gravity Model to enable the estimated trips to match the real number of trips, two constraints are employed in the Gravity Model. Firstly, the total number of departures by a PT mode from a public stop m should be equal to the sum of trips that originate from that particular stop to each possible destination stop n:

$$o^{pt}_{m_i}(t) = \sum_{j=1}^{N} od^{pt}_{m_i,n_j}(t), \tag{3.4}$$

and secondly, the total number of trips taken by PT arriving at a stop n at the time interval t equals the sum of trips that terminate at that particular destination from all possible origin stops m:

$$d^{pt}_{n_j}(t) = \sum_{i=1}^{N} od^{pt}_{m_i,n_j}(t) \tag{3.5}$$

The parameters: The time-dependent weights ($A^{pt}_{m_i}$ and $B^{pt}_{n_j}$) are the parameters of the Gravity Model that are estimated iteratively. The estimation equations can be transformed from Equation 3.3 and Equation 3.4 to:

$$A^{pt}_{m_i}(t) = \frac{1}{\sum_{j=1}^{N} B^{pt}_{n_j}(t) d^{pt}_{n_j}(t) f\left(c^{pt}_{m_i,n_j}(t)\right)} \tag{3.6}$$

and from Equation 3.3 and Equation 3.5 to:

$$B^{pt}_{n_j}(t) = \frac{1}{\sum_{i=1}^{N} A^{pt}_{m_i}(t) o^{pt}_{m_i}(t) f\left(c^{pt}_{m_i,n_j}(t)\right)} \tag{3.7}$$

The criterion: The criterion of the convergence follows either the maximum number of iterations that have been reached:

$$\zeta \in \{1, \ ..., \zeta_{max}\}, \tag{3.8}$$

or the functions of acceptable distance between iterative count ζ and $\zeta + 1$:

$$cc^{pt} \geq \max_{i,j}\left(\max_{i}\left(\frac{A^{pt,\zeta+1}_{m,i} - A^{pt,\zeta}_{m,i}}{A^{pt,\zeta+1}_{m,i}}\right), \max_{i}\left(\frac{B^{pt,\zeta+1}_{n,j} - B^{pt,\zeta}_{n,j}}{B^{pt,\zeta+1}_{n,j}}\right)\right) \tag{3.9}$$

where cc^{pt} is the criteria of convergence that defines the acceptable distance of the last two time-dependent weights, and ζ is the count of iterative calculations. The ζ_{max} used in this research is defined as shown in Assumptions (see Section 3.3.2).

GA vector estimation

The trip generation and attraction (GA) estimation is known as the first step in the traditional four-step demand estimation. This step aims to produce a GA vector that can be used as the total number of departing trips ($d^{pt}_{n_j}$) and total arriving trips ($o^{pt}_{m_i}$) in the Gravity Model-based OD matrix estimation. As introduced in Ortuzar S. and Willumsen, 2011, in practice, the GA vector is initially obtained from the demographic data through a regression analysis. For our research study, the total generation from a stop is the total number of departing trips; thus, the vector of generation is the total number of tap-on at each stop. Meanwhile, the total attraction of a stop is the total number of arriving trips; thus, the vector of attraction is the total number of tap-offs at each stop. Therefore, the GA vector can be captured from historical smart card tap-on/tap-off data as follows:

$$o^{pt}_{m_i}(t) = \sum_{j=1}^{N} od^{pt,historical}_{m_i,n_j}(t), \tag{3.10}$$

$$d^{pt}_{n_j}(t) = \sum_{i=1}^{N} od^{pt,historical}_{m_i,n_j}(t) \tag{3.11}$$

and the GA vector can be represented as $GA^{pt}_{m_i,n_j} = [o^{pt}_{m_i}, d^{pt}_{n_j}]$.

The chapter aims to demonstrate the effectiveness of using the Gravity Model for estimating the PT OD matrix with the most accessible data: total generation and attraction and travel cost data. The data provided for this research study is stop-based smart card data, we process it as stop-based GA vectors. The travel cost data obtained from both traffic states and topological graphs is calibrated following the deterrence function-based Gravity Model by the Traverse searching method (Ortuzar S. and Willumsen, 2011), Hyman's method Hyman, 1969 and the new proposed Entropy-weighted method. We compare the performance of travel cost features following "single travel cost feature", "multiple travel cost features" and the "entropy-weighted multiple travel cost features".

Deterrence Function Considering Single Cost Features

In the OD estimation method via the Gravity Model, the deterrence function, $f\left(c^{pt}_{m_i,n_j}\right)$, is the negative function limiting the number of trips generated from origins to destinations in a network. Following three classical forms of the deterrence function, the friction elements can be described as:

Power function:

$$f\left(c^{pt}_{m_i,n_j}\right) = \gamma(c^{r,pt}_{m_i,n_j})^{\alpha} \tag{3.12}$$

Exponential function:

$$f\left(c^{pt}_{m_i,n_j}\right) = \gamma e^{-\beta c^{r,pt}_{m_i,n_j}} \tag{3.13}$$

Tanner function:

$$f\left(c^{pt}_{m_i,n_j}\right) = \gamma(c^{r,pt}_{m_i,n_j})^{\alpha}\ e^{-\beta c^{r,pt}_{m_i,n_j}} \tag{3.14}$$

where γ is the weight of the travel cost, in the case study of this chapter, $\gamma = 1$, in order to examine the performance of raw travel cost features in OD estimation. $i, j = \{1 \dots J\}, m, n = \{1 \dots N\}$; α and β are parameters of the deterrence function that are required to be estimated.

Deterrence Function Considering Multiple Cost Features

However, in cases where travel costs such as travel time and connection are used in our situation, due to the property of non-linearity with respect to travel distance, a simple summation, as employed in Rubio-Herrero and Muñuzuri, 2021, is not suitable. Therefore, after fitting using a deterrence function(Equation 3.12, Equation 3.13 or Equation 3.14), we standardise the travel cost matrix by dividing it by the mean fitted travel cost ($\overline{f\left(c^{r}_{m_i,n_j}\right)}$). This normalisation ensures that all travel cost features have a consistent mean value of 1. Consequently, we can aggregate all travel cost matrices to create a fused travel cost matrix that encompasses the effects of multiple types of travel costs. The fusion travel cost can be represented by:

$$f\left(c^{pt}_{m_i,n_j}\right) = \sum_{i=1}^{R} \frac{f\left(c^{r}_{m_i,n_j}\right)}{\overline{\overline{f\left(c^{r}_{m_i,n_j}\right)}}}. \tag{3.15}$$

Deterrence function calibration

The calibration of the deterrence function is significant as it directly impacts the subsequent performance of OD estimation. In our effort to verify this idea and determine the most suitable calibration approach for our specific case study, we conduct the calibration analysis involving the Hyman's method.

Deterrence function calibration by the Hyman method: The technique introduced by Hyman has proven to be effective due to its fast convergence. Such a method initialises the parameter by:

$$\beta_0 = \frac{3}{2\bar{C}}, \tag{3.16}$$

where β_0 is the parameter of the deterrence function and $\bar{C}$ is the mean travel cost. In a new cycle of calibration,

$$\beta_\zeta = \frac{\beta_0 c_0}{\bar{C}}, \tag{3.17}$$

where c_0 is the mean estimated travel cost. For the following run of the calibration with $\zeta \geqslant 1$,

$$\beta_{\zeta+1} = \frac{(\bar{C} - \bar{C}_{\zeta-1})\beta_\zeta - (\bar{C} - \bar{C}_\zeta)\beta_{\zeta-1}}{\bar{C}_\zeta - \bar{C}_{\zeta-1}}, \tag{3.18}$$

where c_0 is the mean estimated travel cost. Such an estimation method is investigated in this chapter for comparison purposes, as well.

Following the Hyman method, we engage in an iterative adjustment of the parameters α and β in three common forms of the deterrence function: power (request a calibration on α in Equation 3.12), exponential (request a calibration on β in Equation 3.13), and Tanner (request a calibration on both α and β in Equation 3.14). The calibrated deterrence function, which accurately models travel cost and influences trip distribution, will be utilised in Equation 3.3 for iteratively estimating the OD matrix until the specified objective is achieved (as defined in Equation 3.8 and Equation 3.9).

According to Hyman, 1969 and Williams, 1976, the solution of Equation 3.4 and Equation 3.5 follows an iterative calculation process. Since the mean observed travel cost should equal the estimated one, we introduce another constraint:

$$\sum_{i,j}^{J} od_{i,j} c_{i,j} = \sum_{i,j}^{J} od^*_{i,j} c_{i,j}. \tag{3.19}$$

where the exponential function is given as the deterrence function. For the power function, the constraint should be defined as:

$$\sum_{i,j}^{J} od_{i,j} \log c_{i,j} = \sum_{i,j}^{J} od^{*}_{i,j} \log c_{i,j}, \tag{3.20}$$

for Tanner function, both Equation 3.19 and Equation 3.20 need to be considered.

To adjust the travel cost matrix using the deterrence function, it requires an initial β_0; according to Hyman, 1969,

Step 0:

$$\beta^0 = \frac{3}{2\overline{f(c)}} \tag{3.21}$$

where the mean observed travel cost $\overline{f(c)}$ is defined as:

$$\overline{C} = \sum_{i,j}^{J} c_{i,j}. \tag{3.22}$$

Step 1: Therefore, the initial friction factor matrix can be converted from the cost matrix to Equation 3.12, Equation 3.13 or Equation 3.14. According to Ortuzar S. and Willumsen, 2011, the initial trip matrix can be obtained by:

$$od^0_{i,j} = \frac{\sum_{i,j}^{J} f((c_{i,j}))o_i}{\sum_{i,j}^{J} f((c_{i,j})}, \tag{3.23}$$

and the time-dependent weights ($A^{pt}_{m_i}$ in Equation 3.6 and $B^{pt}_{n_j}$) in Equation 3.7 can be updated accordingly.

Step 2: With the base matrix, we can update the initial mean estimated travel cost, following the method proposed by Williams, 1976:

$$C^0 = \frac{\sum_{i,j}^{J} f((c_{i,j}))o_i d_j}{\sum_{i,j}^{J} f((od_{i,j})^2}, \tag{3.24}$$

and the mean observed travel cost becomes:

$$\overline{C^s} = \frac{\sum_{i,j}^{J} od^s_{i,j} c_{i,j}}{\sum_{i,j}^{J} od_{i,j}}, \tag{3.25}$$

where s represents the iterative number in gravity estimation.

Step 3: According to Hyman, 1969, the parameter for the deterrence function can be computed by:

$$\beta^1 = \frac{\overline{C^0} * \beta^0}{\overline{C}^0}, \tag{3.26}$$

then the cost matrix can be re-adjust using the β^1, returning to *Step 1*. However, for $s \geq 1$, the discrepancy of deterrence factor between two iterations should be proportional to that of mean travel costs to avoid calculation error according to Williams, 1976; therefore:

$$\beta^{s+1} = \beta^{s} + \frac{(\bar{C} - C^{s})(\beta^{s} - \beta^{s-1})}{C^{s} - C^{s-1}}. \tag{3.27}$$

Deterrence function calibration by the Traverse Searching method: The decision to employ the traverse method for determining optimal parameters in research depends on the complexity of the problem and the characteristics of the dataset (Suprayitno, 2018). In light of our experience and the outcomes of applying the Hyman method, we observed that the parameters' impact on the OD estimation can be unpredictable. Hence, we conducted a traverse sensitivity test to explore the optimal parameter that best estimates the OD matrix. The range of parameters used in this research is defined in Assumptions (see Section 3.3.2).

Based on the outcomes derived from Hyman's methodology, we confirm the most suitable form of the deterrence function for fitting purposes. This selected function is then utilised for Traverse Searching. Subsequently, guided by the suggested parameters of the deterrence function following Hyman's method, we define a range of potential parameters for exploration. Within this range, our objective is to identify the optimal set of parameters that enhance estimation accuracy to a greater extent.

Entropy-weighted Ensemble Cost Features

Assigning weights to travel cost features ensures their appropriate impact on OD estimation, thereby enhancing the accuracy of the estimation. However, weighting requires considerable information, which has historically posed a challenge. Our results from testing different travel cost features individually in the deterrence function (as discussed in Section 3.2.2) reveal that various travel costs perform differently when fitted to the same deterrence functions, which underscores the significance of selecting suitable travel cost features in OD estimation.

Our proposed method is the initiative to consider weighing various travel cost features to maximise the accuracy of OD estimation. Such a feature resemble method is derived from the principles of Shannon's entropy Shannon, 1948; McClean, 2003 which has been used for feature ranking (see Ai, 2017; Qi et al., 2021; Wei et al., 2022; Nie et al., 2016). In this way, our current research study takes a further step and applies entropy when weighing various travel costs and evaluating their true importance for the final trip demand estimation process.
In the following, we detail our proposed weighting method via Shannon's entropy:

Network representation: The transport network is captured by an unweighted non-directed graph which we denote $G = (V, E)$, and which follows the Space L' representation. The set of vertices is represented by $V(G) = \{v_1, v_2, \dots, v_N\}$, where N is the total number of PT stops, while the elements of E are the edges following $E(G) = \{e_1, e_2, \dots, e_l\}$. For a PT mode pt, let G^{pt} be the graph where V^{pt} is the set of vertices, and E^{pt} is the set of edges.

Travel cost feature representation: Each network has its unique topological feature reflected by centrality and global characteristics (see Lin and Ban, 2013). Currently, the most used centrality measures in the literature are connection, closeness and network straightness. The details of these topological features are described in Section 3.2.2. In this research, the travel cost feature representations also include the traffic features such as the fare costs, the travel distance and time. Therefore, the travel cost features of the graph G include both topological and traffic features, which are further represented by $C(G) = \{c_1, c_2, \dots, c_R\}$, where the total number of travel cost features is R.

According to the graph features, the matrix following travel cost features by nodes ($c_{m,r}$) can be expressed as:

$$S = \begin{bmatrix} s(v_1, c_1) & s(v_1, c_2) & \dots & s(v_1, c_R) \\ \dots & \dots & \dots & \dots \\ s(v_N, c_1) & s(v_N, c_2) & \dots & s(v_N, c_R) \end{bmatrix} \tag{3.28}$$

where for each value of the cell, s is the travel cost value defined by the location, which is the PT stop v (ordered as $1 \dots N$), and the travel cost feature c, either topological or traffic features (represented by $1 \dots R$). In this way, for each row in this matrix, the row number represents the stop name; for each column, the column name represents either topological or traffic features. Therefore, the $s(v_1, c_1)$ means the value of the first travel cost feature for the first stop, and $s(v_1, c_2)$ is the value of the second type of travel cost feature for the first stop. To simplify computation, we convert origin-destination travel costs (c_{mn}), like travel time, into one-dimensional factors. Each value is assigned to its origin stop (c_m), aligning its dimension with other one-dimensional features such as closeness. We employ ($s_{m,r}$) to represent the origin stop-based travel cost.

Standardised travel cost matrix: To standardise a feature r for each node (PT stop), the ratio is estimated by using the mathematical formula below:

$$u_{m,r} = \frac{s_{m,r} - min(s_{m,r})}{max(s_{m,r}) - min(s_{m,r})} \tag{3.29}$$

The standardisation technique described is called max-min scaling, also known as max-min normalisation. Through normalisation, each feature contributes to the final distance roughly in proportion to its range. The scaling range following max-min scaling is set to [0, 1].

Thus, the standardised travel cost matrix is denoted as:

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,R} \\ \dots & \dots & \dots & \dots \\ u_{N,1} & u_{N,2} & \dots & u_{N,R} \end{bmatrix} \tag{3.30}$$

Entropy measure: According to Shannon's entropy (see Ai, 2017; Shannon, 1948; McClean, 2003), the ratio of each standardised travel cost feature $u_{m,r}$ is denoted by impact probability $p_{m,r}$, where:

$$p_{m,r} = \frac{u_{m,r}}{\sum_{r=1}^{R} u_{m,r}} \tag{3.31}$$

which helps us to estimate further the entropy of each travel cost measure, which is denoted by:

$$I_r = -\frac{1}{N} \sum_{m=1}^{N} p_{m,r} \log(p_{m,r}). \tag{3.32}$$

We apply the max-min scaling again to standardise the weights, as can be expressed by:

$$w_r = \frac{I_r - min(I_r)}{max(I_r) - min(I_r)}. \tag{3.33}$$

Considering that the weights are entropy-based factors representing uncertainty, where features with higher entropy exhibit greater uncertainty and unpredictability, we aim to prioritize features that are stable and predictable in OD estimation. Therefore, we calculate the importance of each feature as $1 - w_r$. Therefore, according to Equation 3.15, the scaled fusion travel cost features becomes:

$$f\left(c^{pt}_{m_i,n_j}\right) = \sum_{i=1}^{R} \frac{(1-w_r) f\left(c^{r}_{m_i,n_j}\right)}{\overline{(1-w_r) f\left(c^{r}_{m_i,n_j}\right)}}. \tag{3.34}$$

where each importance factor is assigned to the travel cost feature accordingly.

As the weighted travel cost features are constructed, serving as a calibrated friction matrix that can be readily applied in the Gravity Model. Consequently, utilising the Gravity Model to explore the optimal balancing factor that maximises estimation accuracy, with the goal of minimising disparities between the estimated and historical OD matrices, as demonstrated in this chapter, or trip length distributions, as utilised in prior literature.

Topological Features

In the discussion in Section 3.2.2 and Section 3.2.2, we include multiple travel cost features, in this section, we discuss those features in detail. Apart from the connection derived from the adjacency matrix, additional topological cost features captured from the network graph including closeness and straightness are also selected to enhance the accuracy of OD matrix estimation in this study. These features capture attributes related to travel distance that influence passengers' route choices and, consequently, impact trip distribution. The definition of each feature is expressed through the following equations, the same as described in Lin and Ban, 2013:

Connection: property indicates the number of edges connected to a node, also known as degree in graph theory, which is determined based on an adjacency matrix as follows:

$$c^{cn,pt}_{m_i,n_j} = \sum_{m_i=1}^{J} \sum_{m_j=1}^{J} ed_{m_i,n_j}. \tag{3.35}$$

Closeness: the characteristic defining the total travel distance from a given node to all other accessible nodes in the network and is expressed as:

$$c^{cl,pt}_{m_i,n_j} = \frac{1}{\sum_{m_j=1}^{J} d_{m_i,n_j}}, \tag{3.36}$$

where d_{m_i,n_j} indicates the travel distance on predefined bus routes, which is the shortest path travelled by bus.

Straightness: the feature displaying the ratio of the Euclidean distance ($d^{Eucl}_{m_i,n_j}$) over the shortest travel distance following the bus routes.

$$c^{st,pt}_{m_i,n_j} = \sum_{m_j=1}^{J} \frac{d^{Eucl}_{m_i,n_j}}{d_{m_i,n_j}}, \tag{3.37}$$

3.2.3 OD Matrix Evaluation

Assuming that the estimated OD matrix using our proposed approach is denoted as $[O\hat{D}_t]$, while the observed one is $[OD_t]$, then the OD estimation accuracy in this research is measured by using the following three performance metrics:
Mean Absolute Error (MAE):

$$MAE = \frac{1}{N} \sum_{n=1}^{N} \left| [O\hat{D}_t] - [OD_t] \right| \tag{3.38}$$

where n represents the PT stop and N is the total number of stops in the network.
Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left([OD_t] - [O\hat{D}_t] \right)^2} \tag{3.39}$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{[OD_t] - [O\hat{D}_t]}{[O\hat{D}_t]} \right| \times 100\% \tag{3.40}$$

where t is the iteration number and T represents the total number of iterative times.

3.3 Case Study

3.3.1 Geography and Data Information

The zones covered in the study are located to the Northwest of Sydney, along the M2 motorway, which includes several major residential and business areas, as shown in Figure 3.2 upper-left. This area is defined following the digital mapping according to the Statistical Area Level 2 (SA2) Australian Bureau of Statistics, 2021. The zones used by the car network are defined in the software Aimsun according to a Voronoi diagram Xiao et al., 2016; Nikolić and Bierlaire, 2018, which is denoted as $Z_j,\ j \in \{1 \dots J\}$, as shown in the Figure 3.2 (up left). As of 2017, there are 76 bus routes, two train routes and a metro route spread within this area. There are 3799 bus stops and seven train and metro stations, as shown in Figure 3.2 (down). To simplify notations, we will further refer to our case study as the M2 area in the following sections. The smart card data system combines records for both trains

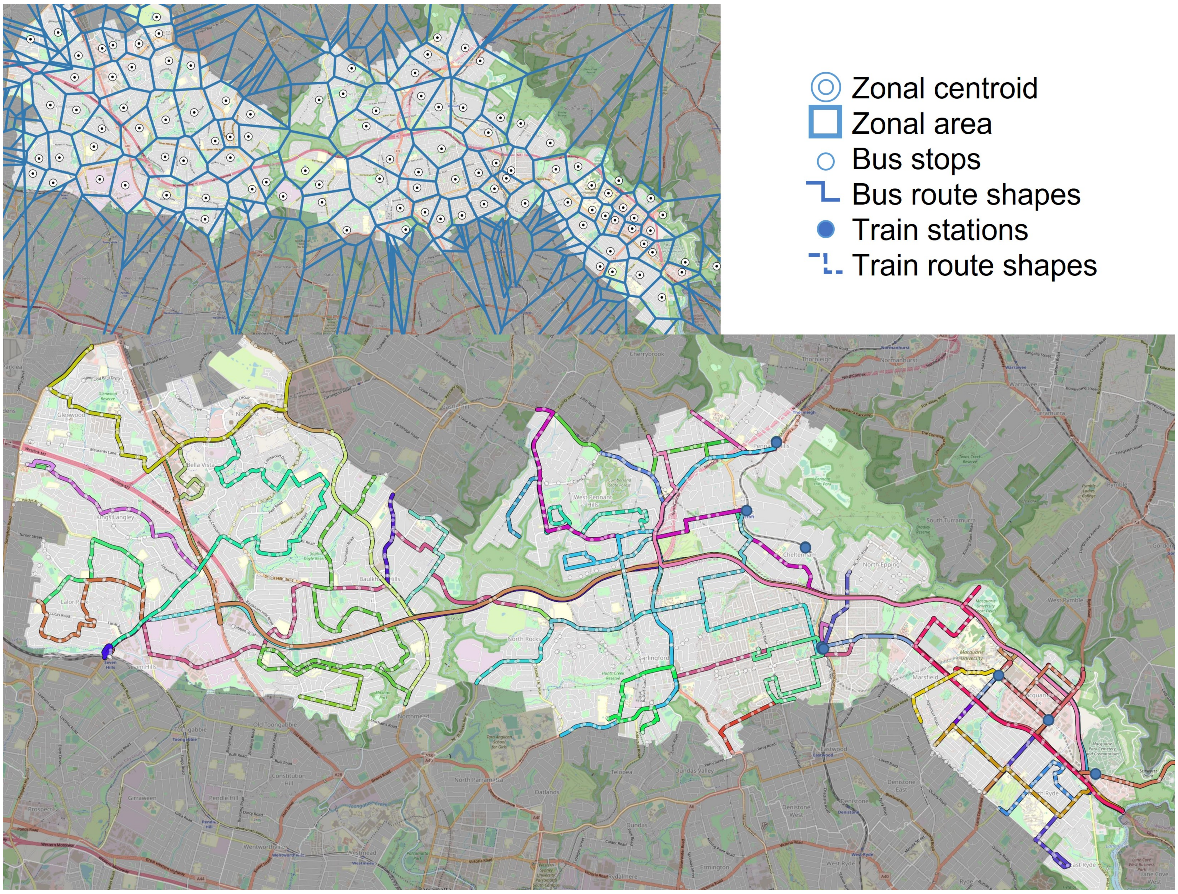


FIGURE 3.2: Zones in M2 area (up left) and the 2017 Sydney M2 area PT, including buses, trains and metro networks (down).

and metros. Consequently, this chapter treats trains and metros as a single mode of transport, referred to as "train", throughout this study.

The trip distribution is captured from the local smart card data. The raw smart card data has been processed and filtered in advance to eliminate outliers and anomalies. The trip distribution in Figure 3.3 is drawn by using as an example one month of smart card data (June of 2017) for the M2 area in Sydney. The upper figure in Figure 3.3 illustrates the distribution for trips by bus, and the rest figure in Figure 3.3 illustrates the distribution of train trips. As shown in the Figure 3.3, for both transport services, the morning peak hour starts from 7:00 and lasts until around 11:00. In contrast, the afternoon peak hour spreads from 16:00 to 20:00. In the case study exemplified in this chapter, we focus mainly on the morning peak hour (as the afternoon can follow a similar approach); therefore the data for 7:00 to 11:00 is collected and used for the estimation method.

The topological feature data for the cost matrix estimation is captured from historical PT GTFS data provided by the OpenData OpenData, 2017 and another open source TransitFeeds TransitFeeds, 2017. The data include the information of the PT agency, its calendar, the routes information, the PT stop times and stop location, as well as all the information regarding the PT trips and stops. The data for June 2017 is collected and used in this section as an exemplification.

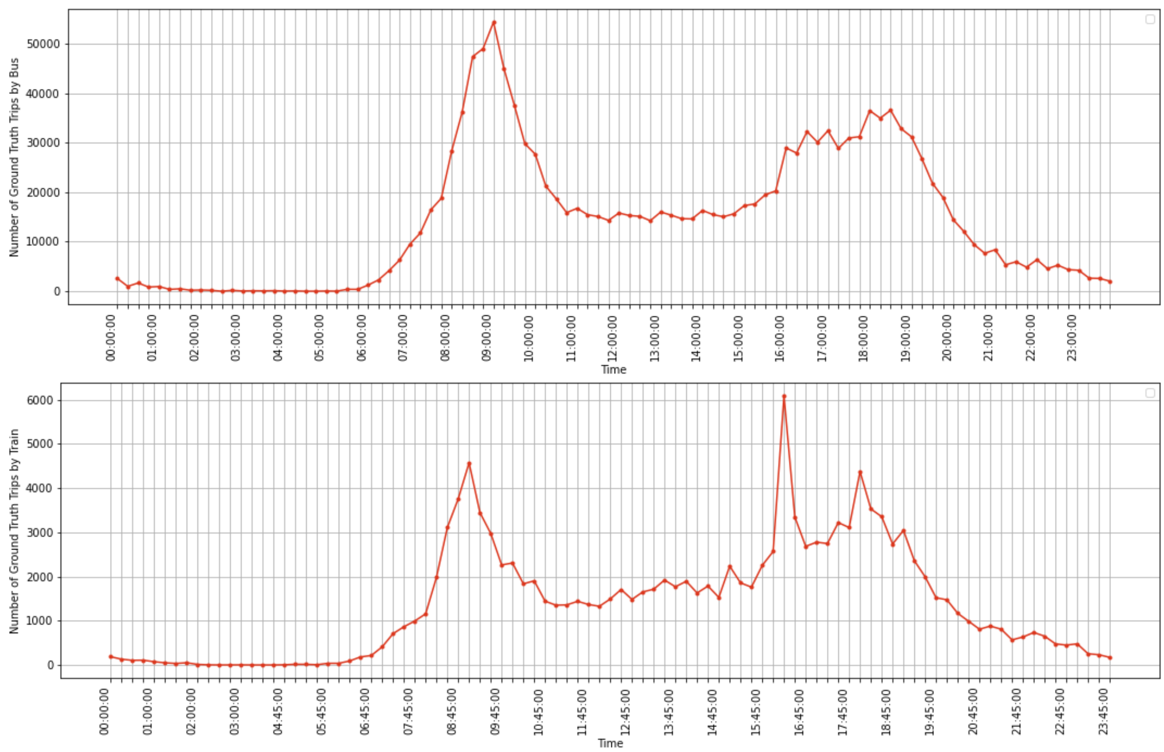


FIGURE 3.3: Ground truth PT trip distributions by time for buses (up) and trains (down).

3.3.2 Assumptions

External nodes: In this research study, an external node is inserted in the OD estimation approach using the Gravity Model detailed in *Stage 3* to balance the total generation and attraction trips. Since the M2 area is only a tiny part of the Great Sydney area, apart from the trips completed within the study area, there are still trips that solely start in the M2 area, which requires a match of an external node to attract those trips. In contrast, for those trips that simply end in the M2 area, an exterior node is also required to generate these trips. This way, we can reach a balanced GA vector that matches the constraints of Equation 3.4 and Equation 3.5.

Convergence criteria: In the case study, the maximum iteration time for each experiment is 20, and the convergence criteria are reached when the gap between estimated and ground-truth total generation and attraction is less than 0.01%. In this research study, we aim to identify the most suitable method for model calibration—one that is accessible, adaptable, and converges quickly. To achieve this, we strive to set the value of r_{max} as high as possible, ensuring that the model satisfies Equation 3.9 while also keeping the computational requirements for iterative calculations manageable on our computer. Consequently, we conduct all experiments until they meet the acceptable criteria multiple times. The maximum number of iterative counts required for convergence in all experiments is 18. Hence, we round up this number to 20 for input as r_{max}.

Hyman method using Tanner function: In the literature, Hyman (Hyman, 1969) demonstrated deterrence function calibration using only the exponential function form, with a single parameter, β, defining the function. Following the same approach, we apply the calibration method to a power function (see Equation 3.12). To streamline calculations and demonstrate the performance of the

Tanner function (refer to Equation 3.14), we utilise identical values for α and β when considering the Tanner function.

Traverse Searching method limitations: When exploring values of α and β using the Traverse Searching method (refer to Section 3.2.2), according to the results following Hyman's calibration method, a range from 0 to 3 in intervals of 0.05 is used for the train network in this chapter. For the bus network, in instances where a distinct pattern was not apparent, we extended the range to encompass values from 0 to 0.2 with intervals of 0.01. For the sake of time and computational considerations, we keep these parameter pairs constant throughout the iterative estimation processes.

3.3.3 Scenario and Experiments Configuration:

When estimating the deterrence function, we set up three main scenarios that match the three cost matrix estimation methods presented in *Stage 2*, namely "Deterrence Function Considering Single Cost Features" *(S1)*, the "Deterrence Function Considering Multiple Cost Features" *(S2)*, and the "Entropy-weighted Ensemble Cost Features" *(S3)* considering different travel cost features covering traffic characteristics including travel time, travel distance and fare cost, as well as graph topological features including connections, closeness and straightness, as shown in Table 3.1

TABLE 3.1: Scenarios and experiments settings.

Network	Calibration Method	Travel cost type	Optimal fitting function	Optimal travel cost	Optimal combination of costs
Bus/train	Traverse searching (S1)	Single cost (S1E1)	Exponential/Power/Tanner	[Connection, closeness, straightness, efficiency, fare cost, travel time, travel distance]	
	Hyman (S2)	Single cost (S2E1)			
		Multiple costs		Fusion Costs (S2E2)	Optimal combination (S2E3)
	Entropy (S3)	Multiple costs		Fusion Costs (S3E1)	Optimal combination (S3E2)

TABLE 3.2: Example of list of travel cost features combinations.

Combination No.	Travel cost features included	Combination No.	Travel cost features included
1	connection,closeness	39	connection,closeness,fare_cost,TravelTime
2	connection,straightness	40	connection,closeness,fare_cost,TravelDistance
3	connection,fare_cost	41	connection,closeness,TravelTime,TravelDistance
4	connection,TravelTime	42	connection,straightness,fare_cost,TravelTime
5	connection,TravelDistance	43	connection,straightness,fare_cost,TravelDistance
...	...	...	...
15	TravelTime,TravelDistance	53	connection,closeness,straightness,TravelTime,TravelDistance
16	connection,closeness,straightness	54	connection,closeness,fare_cost,TravelTime,TravelDistance
17	connection,closeness,fare_cost	55	connection,straightness,fare_cost,TravelTime,TravelDistance
18	connection,closeness,TravelTime	56	closeness,straightness,fare_cost,TravelTime,TravelDistance
19	connection,closeness,TravelDistance	57	connection,closeness,straightness,fare_cost,TravelTime,TravelDistance

In *S1*, all six travel cost features are incorporated into three forms of deterrence functions using the Traverse Searching method, yielding 21 outcomes. The range of potential parameters is determined based on the outcomes using Hyman's method, as detailed in *Traverse Searching method limitations* of Section 3.3.2.

In *S2*, following Hyman's searching method, all six travel cost features are incorporated into three forms of deterrence functions, as well. Under this setup, we incorporate three sub-experiments, each testing the efficacy of single costs, multiple costs, or optimally combined fusion costs. Specifically, the experiments of using single costs *(S2E1)* cover all travel cost features individually, whereas multiple costs *(S2E2)* involve a fusion of these features into a unified representation. Conversely, optimally combined fusion costs *(S2E3)* follow a comparative evaluation of all feasible combinations of diverse travel costs, selecting the most optimal combination for utilisation in OD estimation. Such combination test results in 57 outcomes. An example of the combination list is shown in Table 3.2.

In *S3*, we assess the performance of each combination of travel cost features, resulting in 57 outcomes. This analysis provides evidence regarding the optimal selection of travel cost features that maximise benefits in OD estimation. Both the fusion of all travel cost features *(S3E1)* and the optimal combined fusion costs *(S3E2)* are tested under this setup.

3.4 Results

3.4.1 Correlation across travel cost features

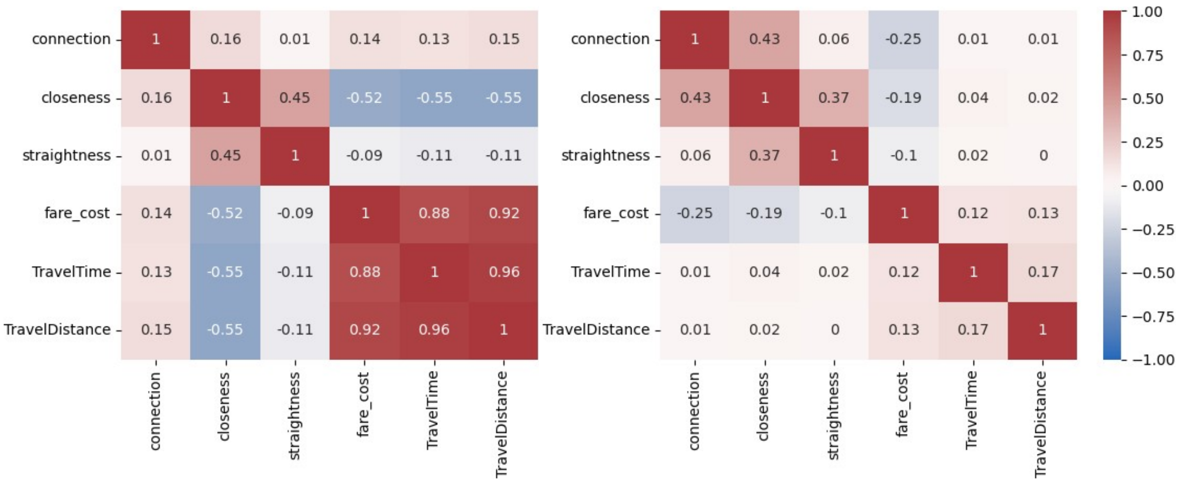


FIGURE 3.4: Correlation matrix across six travel cost features for bus networks (left) and train networks (right).

A correlation matrix helps identify relationships between variables and is useful for understanding patterns and dependencies in the data. In our case study, we consider six travel cost features that have a deterrence effect on trip distribution to estimate the OD matrix for PT networks. Three of these six features are traffic characteristic-related, while the other four features are graph topological features. They are all related to the PT networks, but the internal relationships between each feature are still unknown. We want to identify the overlapping features or any new correlation between features that have never been found. In order to see the correlation between features, we visualise the assessment using a correlation matrix.

In the correlation matrix illustrated in Figure 3.4, each cell represents a Pearson correlation coefficient. A coefficient of 1 indicates a strong positive relationship between variables, 0 indicates a neutral relationship, and -1 suggests a strong negative relationship. The Pearson correlation coefficient quantifies the linear correlation between two sets of feature data.

In the correlation matrix for bus networks, see Figure 3.4 left, we observe a notable correlation among traffic state features, particularly between travel time and route distance, with a correlation coefficient of 0.96. This correlation arises because both variables capture aspects of the spatial relationship between two points within the network.

However, such a strong correlation is not as evident in the train network. This discrepancy can be attributed to the small size of the train network that we have available for this case study. In addition, train networks, unlike bus networks, operate more according to fixed schedules dictated by passenger demand and constrained by infrastructure limitations (such as capacity, speed, maintenance requirement, and crew schedule). These factors lead to variability in travel time relative to geographical factors, thus weakening the correlation between travel time and route distance in train networks.

Regarding the topological features in the bus network, we can observe that closeness (which reflects the reciprocal of the sum of distances from a node to all other nodes in the network), is related to straightness (which reflects the ratio of the Euclidean distance over the shortest path), with a coefficient of 0.45. Similarly, for a train network, the closeness is related to the straightness, with a coefficient of 0.37. One possible explanation for this correlation could be that a well-connected network with multiple routes may allow PT passengers to travel more in shorter distances.

Additionally, in small train networks, closeness shows a higher relation to connection (reflecting the total number of links connected to a given node), where the correlation coefficient is 0.43. This indicates that the train stations with more connections are likely to be centrally located and accessible to other stations, resulting in shorter average distances to reach other destinations in the network. However, bus networks often have a more decentralised structure, with multiple routes and stops dispersed throughout the network. This decentralised nature reduces the correlation between connection and closeness in bus stations, as there may not be a clear hub-and-spoke pattern as observed in the train network.

3.4.2 Calibration results of deterrence function considering single travel cost

Following the Hyman's method S2: In this section, we assess the performance of six distinct travel cost features across three different forms of the deterrence function, as per Hyman's approach within the Gravity Model framework. The results based on MAPE, presented in Table 3.3, indicate that power forms exhibit effectiveness in a large bus network, with average computation times of 138s. The feature of closeness performs superior to all other features. By using the exponential function, closeness and straightness perform well. The straightness suits Tanner's function better than the rest. The analysis of RMSE results reveals more prominent outliers in the estimated matrix when employing the power function. According to the MAE results, the magnitude of errors between the estimated and historical matrices is relatively low when considering a Tanner function. Notably, the Tanner function does not align with the distribution for travel time.

For a small network such as the train network examined in this case study, we note the remarkable effectiveness of all functions, particularly the exponential function, which displays rapid convergence and calculation. Evidently, all function forms prove suitable for estimating OD superior performance, depicted in Table 3.3, connection and travel time demonstrate superior performance when following an exponential function. However, obtaining travel time by stops necessitates the timetable travel time for all connected trips along public transport routes based on GTFS data, which consumes significant computing time compared to the other features, posing a major drawback to its utilisation. Given that the connection feature also exhibits similarly superior performance, both travel time and connection have been selected for further assessment, as illustrated in Table 3.7.

Following Traverse Searching method S1: Based on the results obtained from Hyman's method, we gain insights into the optimal fitting function and the ideal range of parameters. Subsequently, we conduct a thorough investigation into parameter performance through a traverse search. We select the exponential function for fitting purposes for both bus and train networks.

As depicted in Table 3.4, for the bus network, connection and travel time exhibit the most favourable performance, achieving a MAPE of 35.06%. This stands in contrast to the results obtained from Hyman's method, where the best MAPE of 36.20% was achieved by employing a power function considering closeness. This means the adjustment of parameters within the deterrence function proves to enhance the performance of OD estimation. However, it is imperative to acknowledge that Traverse Searching entails a significant drawback in time consumption. Hence, while investigating parameters and deterrence function forms hold merit, weighing the time investment associated with such endeavours is essential.

For train networks, travel time consistently exhibits superior performance, while the connection feature appears to deviate from its optimal parameter during the search process. This underscores the importance of incorporating the Hyman method (or other pre-search method) for preliminary parameter range exploration.

3.4.3 Calibration results of deterrence function considering multiple travel costs

Following the Hyman's method S2: Acknowledging the effectiveness of integrating multiple travel costs in prior OD estimation studies, we introduce a scenario considering multiple travel costs. We standardise the costs by normalising their mean values to streamline the integration process. Thus, all travel cost features share an equal mean value of 1, facilitating fusion. As demonstrated in Table 3.5, the fusion cost notably enhances the accuracy of the OD estimation for the bus network. However, its efficiency is reduced when applied to the train network. This observation highlights the significance of optimal cost selection or appropriately weighing the costs. Notably, when adopting a fusion cost approach, the exponential form emerges as particularly effective for both types of networks, while the power function proves unsuitable for this specific case study.

Following the Entropy-weighted method S3: *S3* follows the Entropy-weighted method outlined in Section 3.2.2, where we determine the weight assigned to each travel cost feature and construct an ensemble cost for fitting to deterrence functions. The results presented in Table 3.5 reveal that entropy ranking is effective for enhancing OD estimation accuracy in small train networks, facilitating convergence and improvement. However, in the case of bus networks examined in this study, while Hyman's method already yields favourable outcomes, further enhancements achieved through entropy weighting are comparatively modest.

TABLE 3.3: Calibration results of deterrence function considering single travel cost using Hyman's method.

Unit of computing time is second (s)

Bus network							Train network						
Exponential	MAPE	MAE	RMSE	alpha	beta	Computing time	Exponential	MAPE	MAE	RMSE	alpha	beta	Computing time
Connection	38.50%	0.401	27.744	NaN	0.0038	128	Connection	0.83%	0.334	0.638	NaN	0.07	0.56
Closeness	36.66%	0.401	27.892	NaN	0.0325	149	Closeness	1.19%	0.394	0.719	NaN	3.07	0.62
Straightness	36.72%	0.401	27.886	NaN	0.0325	147	Straightness	0.93%	0.362	0.661	NaN	0.07	0.56
FareCost	38.50%	0.401	27.744	NaN	0.0038	129	FareCost	0.88%	0.354	0.643	NaN	0.01	0.58
TravelTime	38.50%	0.401	27.744	NaN	0.0038	128	TravelTime	0.74%	0.329	0.642	NaN	0.38	0.58
TravelDistance	38.50%	0.401	27.745	NaN	0.0038	129	TravelDistance	0.87%	0.353	0.634	NaN	0.01	0.61
Power	MAPE	MAE	RMSE	alpha	beta	Computing time	Power	MAPE	MAE	RMSE	alpha	beta	Computing time
Connection	36.59%	0.401	28.194	0.0250	NaN	136	Connection	0.92%	0.366	0.659	0.00	NaN	0.66
Closeness	36.20%	0.402	28.285	0.0331	NaN	134	Closeness	26.24%	0.531	1.079	2.17	NaN	0.61
Straightness	36.48%	0.401	28.218	0.0019	NaN	137	Straightness	0.92%	0.374	0.660	0.07	NaN	0.59
FareCost	36.25%	0.402	28.297	0.0669	NaN	138	FareCost	0.92%	0.367	0.660	0.00	NaN	0.58
TravelTime	36.97%	0.401	28.127	0.1963	NaN	140	TravelTime	0.93%	0.367	0.663	0.00	NaN	0.65
TravelDistance	36.59%	0.401	28.194	0.0300	NaN	151	TravelDistance	0.93%	0.367	0.661	0.00	NaN	0.62
Tanner	MAPE	MAE	RMSE	alpha	beta	Computing time	Tanner	MAPE	MAE	RMSE	alpha	beta	Computing time
Connection	38.49%	0.401	27.744	0.0038	0.0038	167	Connection	0.84%	0.333	0.636	0.07	0.07	0.63
Closeness	36.74%	0.400	27.871	0.0294	0.0294	204	Closeness	26.22%	0.531	1.078	2.17	2.17	0.69
Straightness	36.51%	0.400	27.884	0.0300	0.0300	201	Straightness	0.92%	0.368	0.659	0.07	0.07	0.63
FareCost	38.49%	0.401	27.744	0.0038	0.0038	154	FareCost	0.89%	0.355	0.644	0.01	0.01	0.65
TravelTime	NaN	NaN	NaN	NaN	NaN	NaN	TravelTime	0.81%	0.336	0.651	0.35	0.35	0.64
TravelDistance	38.49%	0.401	27.744	0.0038	0.0038	149	TravelDistance	0.86%	0.351	0.632	0.01	0.01	0.70

TABLE 3.4: Calibration results of deterrence function considering single travel cost using Traverse Searching method.

Unit of computing time is second (s)

Bus network							Train network						
Exponential	MAPE	MAE	RMSE	alpha	beta	Computing time	Exponential	MAPE	MAE	RMSE	alpha	beta	Computing time
Connection	35.06%	0.401	28.532	NaN	0.04	8594	Connection	2.30%	0.445	0.872	NaN	0.20	813
Closeness	35.20%	0.402	28.539	NaN	0.03	7486	Closeness	2.29%	0.461	0.850	NaN	2.15	552
Straightness	35.20%	0.402	28.539	NaN	0.04	7712	Straightness	2.29%	0.461	0.850	NaN	1.50	544
Efficiency	35.20%	0.402	28.539	NaN	0.01	8599	Efficiency	2.29%	0.461	0.850	NaN	2.90	540
FareCost	35.08%	0.401	28.532	NaN	0.04	8243	FareCost	6.92%	0.512	1.006	NaN	0.10	985
TravelTime	35.06%	0.401	28.532	NaN	0.04	7199	TravelTime	1.87%	0.440	0.861	NaN	0.15	754
TravelDistance	35.10%	0.401	28.533	NaN	0.04	7940	TravelDistance	17.60%	0.544	1.052	NaN	0.05	3819

TABLE 3.5: Calibration results of deterrence function considering multiple travel costs using Hyman's and Entropy method.

Unit of computing time is second (s)

Bus network							Train network						
Deterrence function considering multiple costs													
Hyman method	MAPE	MAE	RMSE	alpha	beta	Computing time	Hyman method	MAPE	MAE	RMSE	alpha	beta	Computing time
Exponential	20.27%	0.534	4.698	NaN	-13.06	190	Exponential	5.88%	0.432	0.776	NaN	6.42	0.57
Power	NaN	NaN	NaN	NaN	NaN	138	Power	27.88%	0.279	0.555	-22142.24	NaN	0.53
Tanner	23.16%	0.637	6.448	-10.58	-10.58	161	Tanner	31.07%	0.477	0.882	38.02	38.02	0.61
Entropy-weighted deterrence function considering multiple costs													
Entropy method	MAPE	MAE	RMSE	alpha	beta	Computing time	Entropy method	MAPE	MAE	RMSE	alpha	beta	Computing time
Exponential	20.58%	0.491	4.085	NaN	-13.74	198	Exponential	0.98%	0.377	0.670	NaN	0.17	0.58
Power	NaN	NaN	NaN	NaN	NaN	136	Power	38.47%	0.406	0.748	-8.40E+07	NaN	0.26
Tanner	22.65%	0.600	4.817	-3.78	-3.78	206	Tanner	2.90%	0.426	0.762	1.05	1.05	0.66

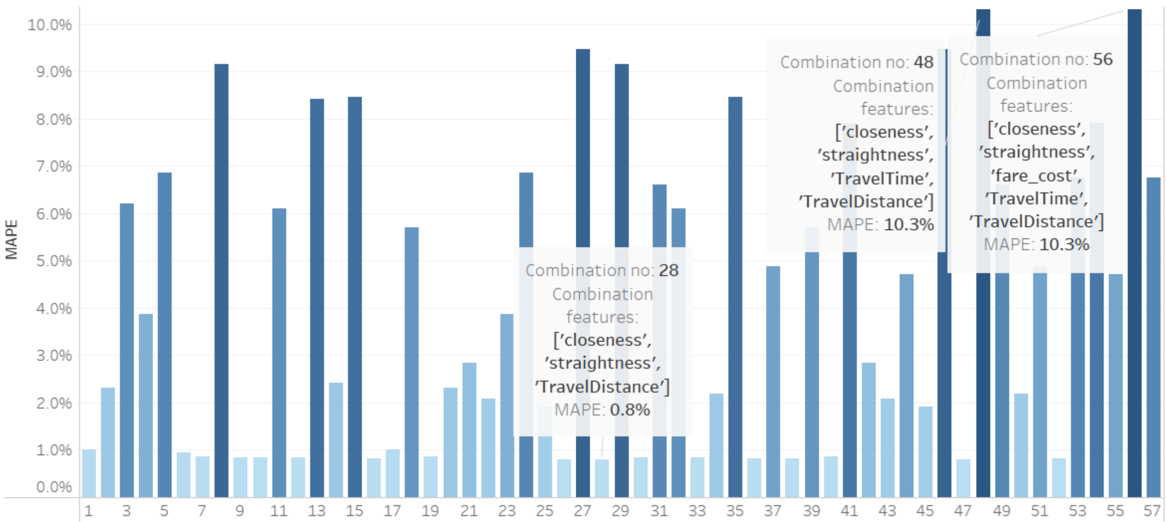


FIGURE 3.5: MAPE for the combination of travel cost features in train networks, fitted by an exponential function using the Entropy-weighted method.

3.4.4 Calibration results of deterrence function considering optimal combination of travel costs

Determination of optimal combination of travel cost features: To further investigate the performance of various travel cost features and their combinations on OD matrix estimation, we conducted a performance comparison of various combinations of travel cost features, as shown in Table 3.2. Such an experiment transversely estimates the OD matrix using all possible combinations of travel cost features based on the Gravity Model.

Due to space constraints, we solely present the comparison results for the train network utilising the Entropy-weighted approach and fitting by the exponential function, and their performance is assessed based on the MAPE, as shown in Figure 3.5. The bars in the figure are arranged in ascending order based on combination numbers ranging from 1 to 57. Notably, the most favourable outcomes are achieved when utilising a combination of closeness, straightness and travel distance (Combination number 28), see evaluation results which are summarised in Table 3.6. The performance of combined travel cost features exhibits variability across different combinations, incorporating additional travel cost features does not always improve the accuracy. Notably, outcomes are less favourable when considering travel time, as evidenced by Combination numbers 48 and 56, which yield comparatively poorer results.

In the case of train networks, when relying on Hyman's method for calibration, the optimal feature combination comprises closeness, straightness and travel distance (Combination number 28) and a combination of closeness, straightness, fare cost and travel distance (Combination number 47), where both combinations achieve a notably low MAPE value of 0.8%. As Combination 47 involves four features, we opt for Combination 28, which comprises three features, as the optimal choice. Evaluation results for this optimal fusion are presented in Table 3.6.

For the bus network, under both the deterrence function utilising Hyman's method and the Entropy-weighted approach, the optimal combinations of features consist of closeness and straightness (Combination number 6). This combination demonstrates minimal MAPE values, indicating their superior

TABLE 3.6: Calibration results of deterrence function considering optimal combination of travel costs using Hyman's and Entropy method.

Unit of computing time is second (s)

Bus network							Train network						
Deterrence function considering optimal combination of costs													
Closeness and Straightness							Closeness, Straightness and Fare Cost						
Hyman method	MAPE	MAE	RMSE	alpha	beta	Computing time	Hyman method	MAPE	MAE	RMSE	alpha	beta	Computing time
Exponential	1.50%	0.015	0.335	NaN	0.7519	128	Exponential	0.82%	0.374	0.669	NaN	0.05	0.66
Entropy-weighted deterrence function considering optimal combination of costs													
Closeness and Straightness							Closeness, Straightness and Travel Distance						
Entropy method	MAPE	MAE	RMSE	alpha	beta	Computing time	Entropy method	MAPE	MAE	RMSE	alpha	beta	Computing time
Exponential	1.50%	0.015	0.335	NaN	0.77	127	Exponential	0.80%	0.366	0.650	NaN	0.21	0.60

performance, with an error rate as low as 1.50% (see Table 3.6). The consistent MAPE values across these combinations suggest that both the Hyman and Entropy methods facilitate effective calibration of travel costs, resulting in optimal outcomes as long as the best feature combination is found.

3.4.5 Comprehensive comparison of calibration methods performance on OD estimation

TABLE 3.7: Summary of optimal calibration method with suitable travel cost features or combined travel cost features.

Network	Calibration Method	Travel cost type	Optimal fitting function	Optimal travel cost	Optimal combination of costs
Bus	Traverse searching	Single cost	Exponential	Connection (B_S1E1)	
	Hyman	Single cost	Power	Closeness (B_S2E1)	
		Multiple costs	Exponential	Fusion Costs (B_S2E2)	Closeness, Straightness (B_S2E3)
	Entropy	Multiple costs	Exponential	Fusion Costs (B_S3E1)	Closeness, Straightness (B_S3E2)
Train	Traverse searching	Single cost	Exponential	TravelTime (T_S1E1)	
	Hyman	Single cost	Exponential	TravelTime (T_S2E1)	
		Single cost	Exponential	Connection (T_S2E2)	
		Multiple costs	Exponential	Fusion Costs (T_S2E3)	Closeness, Straightness and Fare Cost (T_S2E4)
	Entropy	Multiple costs	Exponential	Fusion Costs (T_S3E1)	Closeness, Straightness and Travel Distance(T_S3E2)

Drawing from the analyses conducted above, we consolidate the optimal deterrence function calibration function and features in Table 3.7. This section overviews each experiment's performance metrics, including MAE, MAPE, and RMSE.

Comparison of performance using various calibration methods when optimal travel costs are adopted for bus network: By comparing the bar charts in Figure 3.6, we observe that the method of *S2E3* and *S3E2* performs the best among others in an accurate OD estimation when estimating the OD matrix for the bus network. Their MAE value is 0.015 (see Figure 3.6 a)), which means that by using the proposed method, we can obtain an estimated OD matrix that is approximately 97% more accurate than that estimated by using other methods.

Examining the MAE, it is observed that the inclusion of fusion travel costs leads to increases in the magnitude of errors. However, when considering the MAPE and RMSE, Figure 3.6 b) and c), the utilisation of fusion travel costs demonstrates superior performance by enhancing relative accuracy and effectively reducing large outliers. Implementing the optimal combined travel costs further amplifies

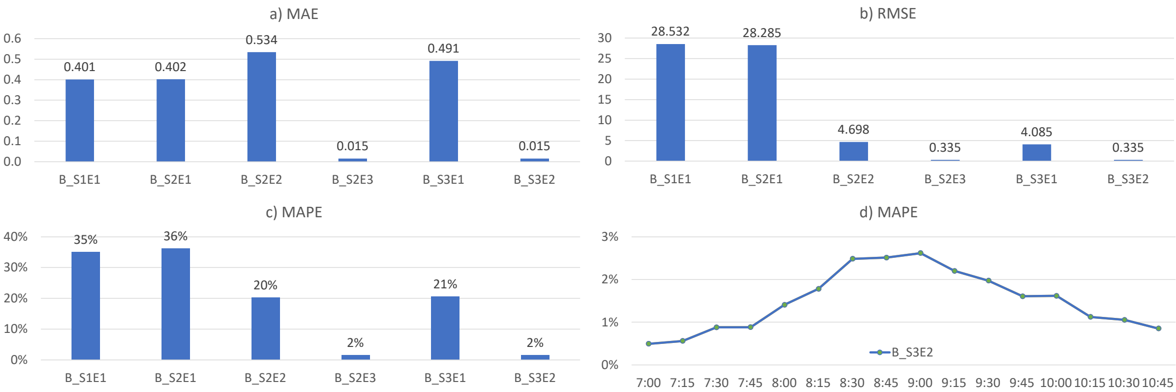


FIGURE 3.6: Bus network errors evaluated between the ground truth and various cost matrix estimations by using: a) MAE b) RMSE c) MAPE, and time-dependent error distribution by using d) MAPE for the proposed Entropy-weighted method.

these benefits, resulting in a 98% reduction in MAPE error for *S2E3* and *S3E2* compared to other scenarios, along with a 95%reduction in RMSE error.

Figure 3.6 d) presents the time-dependent error distribution, showcasing the MAPE for the proposed Entropy-weighted method. The figure illustrates consistently low MAPE values across all time slots. However, a slight peak is evident from 8:30 to 9:30, aligning with the peak patronage hour.

Comparison of performance using various calibration methods when optimal travel costs are adopted for train network: In this comparison, single travel costs (specifically, travel time *S2E1* and connection *S2E2*) outperform the utilisation of multiple travel costs within the proposed method, as indicated by lower MAE and RMSE values. Apart from *S2E1*, where travel time is incorporated into the deterrence function, the performance of the proposed *Entropy-weighted* method stands out according to MAPE, which, on average, is 61.04% less than that for others. We extend our analysis to compare the time-dependent performance of *S2E1*, *S2E4*, and *S3E2*, as depicted in Figure 3.7 (d). Notably, all methods demonstrate consistently low MAPE values across all time slots, unaffected even during peak hours, contrasting with the results observed for large bus networks.

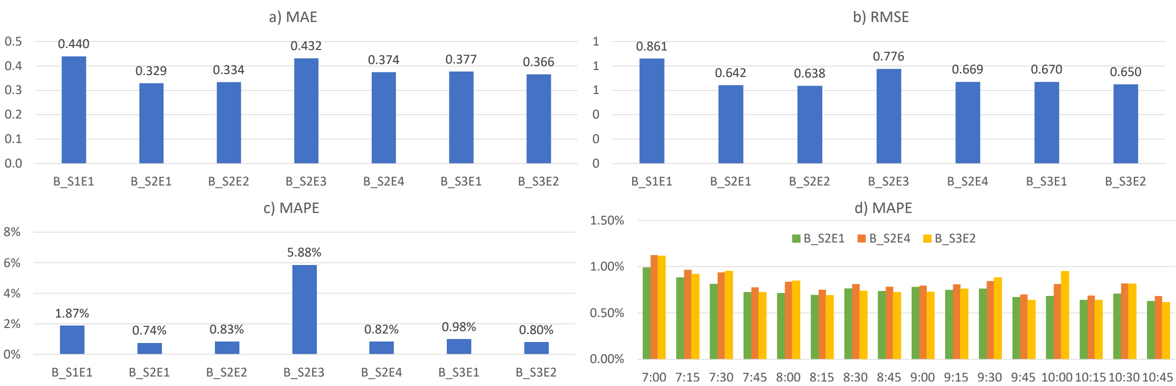


FIGURE 3.7: Trains network errors evaluated between the ground truth and various cost matrix estimations by using: a) MAE b) RMSE c) MAPE, and time-dependent error distribution by using d) MAPE for the proposed Entropy-weighted method.

3.4.6 Importance of travel cost features in demand estimation

We introduced Shannon's entropy-ranking method to evaluate the weights for certain travel cost features. As detailed in Section 3.2.2, each travel cost feature is ranked by Shannon's entropy derived from the impact probability of the cost. We then treat the entropy as a measure of the importance of assessing different travel cost features in estimating the demand matrix. The features with higher impact probabilities and lower entropy scores would be more significant in determining demand, and their inclusion would lead to more accurate demand matrix estimations. As shown in Figure 3.8 below, we display the entropy of travel cost features.

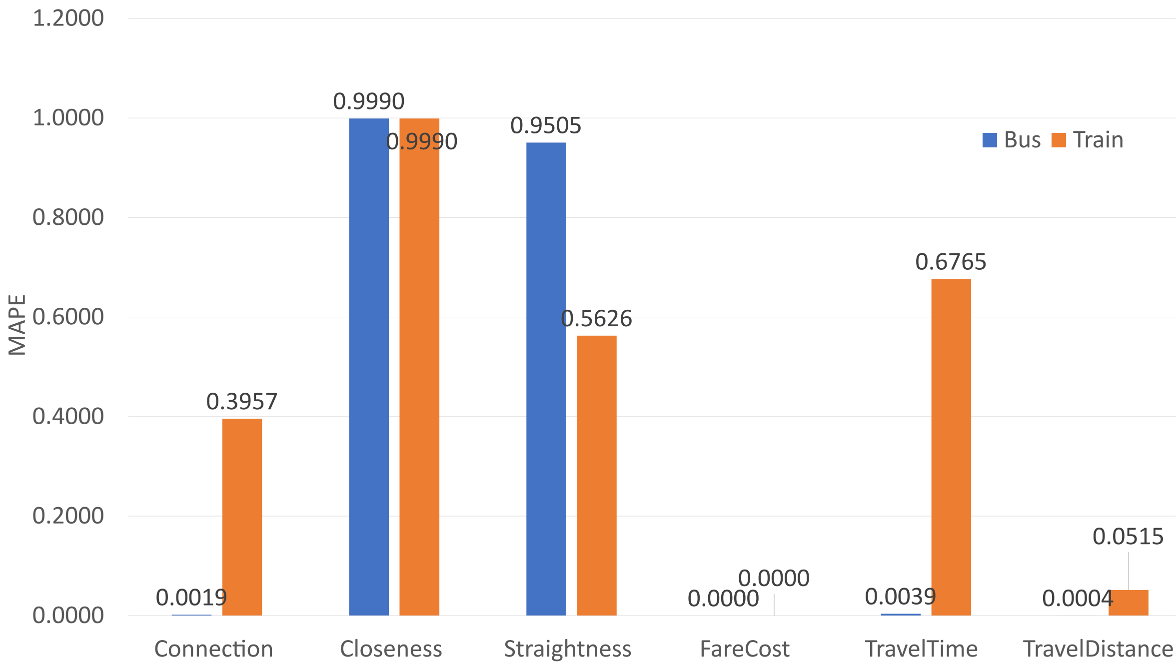


FIGURE 3.8: Normalised importance of each travel cost feature in bus and train OD estimation based on Entropy-weighted method.

The importance of each travel cost feature depends on the entropy calculated based on Shannon's entropy. The entropy calculated for each feature is then processed to determine the percentage of the total entropy contributing to the overall cost. This percentage can be used to determine the relative importance of each feature in influencing the overall cost. Since entropy is a measure of uncertainty, at this stage, the importance determined by the entropy means the uncertainty of the cost feature adds to the overall cost when estimating the demand pattern. Thus, the larger the importance calculated, the higher the influence on the demand estimation.

As shown in Figure 3.8, blue bars represent the importance of each travel cost feature for buses captured by entropy. It is observed that closeness and straightness play a significant role in estimating demand patterns for buses. Network closeness, as introduced in Section 3.2.2, is determined by the connectivity between each node in the network. Therefore, a more connected network is likely to have a higher demand as more effective travel is allowed. The straightness refers to the effectiveness of travelling between each node pair; High straightness in a bus network indicates that routes between stops are relatively direct and straightforward. The high importance of these features matches that passengers are often more willing to use a bus system that provides a direct and short route between their OD. This is the major reason for including a bus system citywide: provide direct and short trips, improve accessibility and offer reliable and affordable transport for more residents.

The orange bars depict the significance of each travel cost feature for trains as determined by entropy analysis. In a small network such as a train network, a broader array of travel cost features influences OD estimation, with closeness, straightness and travel time emerging as the top three influential factors. In contrast to bus networks, the stability of travel time in train networks is enhanced by limited stops, designed schedules, pre-defined routes and adjusted travel speed. Consequently, travel time becomes an essential factor to consider when estimating the OD matrix for train networks.

3.5 Conclusion

This chapter provides a new framework for a dynamic large-scale stop-by-stop OD estimation model for PT. In this model, we emphasise a microscopic stop-based OD matrix, yet in order to simplify the computing workloads, we assume that the time interval is 15 minutes and calculate the number of trips between any OD pair for every 15 minutes to mimic the dynamic condition. The proposed framework shows the ability of our model to examine the effects of the travel cost matrix, namely the "single travel cost feature", "multiple travel cost features" and the "entropy-weighted multiple travel cost features", reflected by various performance metrics (MAE, RMSE and MAPE) between the ground truth matrix and the estimated matrix. The proposed large-scale OD estimation model is established based on the Gravity Model with total generation and attraction (GA vectors) by stops, network physical configuration data and transport services operation data on inputs. In terms of the input data, the smart card data that enables the GA vectors' estimation is used; and the PT GTFS data is processed for the cost matrix data from topological-level cost features, including the connection, closeness, straightness and efficiency, and the traffic-level cost features, such as the travel distance or the travel time as well as the travel distance-based fare costs.

This research study proposed a novel deterrence function calibration method, where Shannon's entropy is employed for weighting the feature for each node (represented by a PT stop). These are due to the fact that the method has pre-weighted the cost features before combining the impact of the cost feature together and applying them in the process of iterative parameter calibration, and the process of weighing the cost features can be separated from the iterative OD matrix estimation, which reduces the load of iterative computing. The performance of the Entropy-weighted method is contrasted with traditional Hyman's and Traverse Searching methods. In bus networks, the Traverse Searching method demonstrates the capability to identify optimal parameters, consequently enhancing estimation accuracy. Moreover, the fusion of travel cost features achieved through both Hyman's and Entropy-weighted methods contributes to further improvements in accuracy. It is worth noting that an optimal combination of travel costs (closeness and straightness) holds the potential to enhance the accuracy of bus OD estimation further. Optimal deterrence function calibration for small train networks is achieved when utilising a single travel cost, such as travel time or connection. While the incorporation of fusion travel costs decreases accuracy in OD estimation, the effective combination of practical travel cost features (closeness, efficiency and fare cost when calibrating by Hyman's method or closeness, straightness and travel distance when calibrating by Entropy-weighting method) leads to a notable increase in accuracy.

Additionally, after illustrating the mean errors by time, the time-dependent tendency of error fluctuation matches the timely number of trips in the network: the peak errors occur when the maximum number of trips occurs in the network (at about 8:30-9:30). This indicates that the network overcrowding is associated with the accuracy of the OD estimation by using the proposed model.

Limitations and Future Directions: In our research study, a framework for a dynamic stop-by-stop OD matrix estimation for large-scale PT is provided. However, the model does not directly include the impact of general traffic, such as the delay time. Although the travel time is captured by using the scheduled time, we could have compared it with the historical travel time captured from smart-card data.

The proposed PT OD matrix estimation model has the potential to be integrated into other transport networks, such as car networks, to form a large-scale multi-modal OD matrix estimation model. In this research, due to the data availability issue, we only establish the model for bus and train networks. The same modelling process can be applied to estimate the OD matrix for cars, light-rails or on-demand solutions.

Furthermore, an area for future research could involve examining transfers between bus and train networks subsequent to OD estimation. Although our historical smart-card data contains limited information on transfers—specifically, we can identify transfer trips but lack details on matching upstream and downstream transfer trips. Subsequent research endeavours may concentrate on aligning upstream and downstream transfer trips and exploring the number of transfers.

In regard to the deterrence function calibration, there is room for further exploration of parameters. However, due to time constraints in this chapter, we have limited our focus to the Hyman and Traverse Searching methods. Many other estimation models, such as the Competing Destination Model (see Thorsen and Gitlesen, 1998), the Self Deterrence Model with Quadratic Cost (SDMQC) (see Fang, Science, and 1995, 1995), or a combination of these models (see Grange, Fernández, and Cea, 2010), hold potential for investigation. Beyond the modelling approach, the form of the deterrence function can also be expanded to include functions like Weibull, Box-Cox, or March (see Rubio-Herrero and Muñuzuri, 2021), among others. Using smart-card data, various other methods can be tested to generate OD matrices based on historical data. These methods include regression models, maximum likelihood estimation, Bayesian estimation, and machine learning algorithms.

Regarding the findings from the traverse search results, we have identified a tipping point where the MAE remains constant regardless of changes in parameter values. Investigating the underlying reasons for this phenomenon could be a direction for future research. Additionally, exploring the tipping point when considering travel cost features beyond those examined in this chapter may provide insights into the relationship between travel cost and the accuracy of estimated OD matrices.

Chapter 4

Modelling public transport disruptions and impact by smart-card data

Evaluating disruptions in public transport utilisation is challenging due to often stochastic traveller behaviour and missing data information on affected services. This chapter proposes a new approach for modelling PT patronage and disruption impact using integrated data-driven modelling and the Fourier transform technique. Firstly, using tap-on and off information of smart-card data, we estimate in-vehicle passenger numbers to integrate as well as trips passing through the incident area. Secondly, considering the PT patronage pattern as a periodic function, we employ the Fourier transform to convert it into a sum of simpler trigonometric functions to filter out the one representing common data noise successfully and generate an accurate profile for a typical day. Thirdly, we introduce an enhanced sensitivity test to improve the model's ability to identify the impact of the disruption. Finally, multiple impact measurement methods are compared to capture the disruption impact.

The findings demonstrate the effectiveness of leveraging in-vehicle count to maximise data volume and enhance impact identification. The PT patronage pattern can be effectively modelled using the Fourier transform. The utilisation of the enhanced sensitivity test can effectively filter out unnecessary trigonometric components, resulting in a refined model capable of accurately identifying the impact of disruption.

This chapter is based on an edited edition of the following article: Zhao D, Mihaita AS, Ou Y, Grzybowska H. Modelling public transport disruptions and impact using smart-card data. IEEE Intelligent Transportation Systems Conference 2023.

Contents

4.1 Introduction

4.1.1 Background and motivation

In a multi-modal transport network, a road traffic disruption always extends beyond its own network. These disruptions can affect related bus lines, train services, and, subsequently, the overall transport network. The bus network shares mobility rights with other road transport modes, and is very sensitive to road traffic disruptions. If an incident occurs on public transport (PT), passengers who are affected by the disruption may seek alternative modes of transportation to continue their journeys. This increased demand for alternative transport services can create unexpected changes in the transportation system. When disruptions occur on major transportation modes like trains, the consequences can be significant. For example, in the case of a train strike in London (see Saberi et al., 2018), the disruption in the train network can lead to city-wide congestion and have a widespread impact on commuters. With limited or no train services available, hundreds and thousands of passengers are affected, and the demand for alternative travel modes, such as shared bicycles, buses or subway services, increases dramatically (see Yuming Ou and Adriana-Simona Mihăiţă and Fang Chen, 2022, Ou, Mihaita, and Chen, 2020).

These unexpected changes in demand can strain the transport network, especially if it exceeds the network's capacity to handle such disruptions. However, it is also essential to consider scenarios where the impact is relatively small, such as a localized disruption caused by a road accident or a vehicle breakdown. These situations are more common in daily life and may not have a city-wide influence, but they could be one of the influence factors for unstable traffic (see Zhao, 2019). This chapter aims to tackle this issue by examining the daily PT patronage profile and investigating the impact of road incidents on this patronage pattern. By understanding the daily movement of PT users in the network, we gain insights for making informed decisions to enhance the PT service and ease the cooperation between private vehicles and PT.

Analyzing the patronage pattern can guide improvements in the supply of PT services, ensuring they align with the needs and preferences of passengers. However, when compared to the travel pattern analysis for private vehicles, analysis regarding PT travel patterns has received relatively less attention in the literature due to limited data availability and modelling methods. To model the PT travel behaviour, data collection can be carried out through various means, such as census or surveys (see Rahimi et al., 2021), smart-card data (see Assemi et al., 2020), or utilizing the General Transit Feed Specification (GTFS) data (see Zhao, 2019). Data-driven methods often rely on machine learning and deep learning algorithms (see Saleh, Grigorev, and Mihaita, 2022, Grigorev et al., 2022b, Grigorev et al., 2022a, Wen et al., 2018, Grigorev et al., 2022b). These techniques utilize the available data to extract meaningful insights and make accurate impact predictions (see Ou, Mihaita, and Chen, 2020; Shafiei et al., 2021b; Wen et al., 2018), and some of them will focus on also predicting how long disruptions will last in the network Mihaita et al., 2019, Shafiei, S. and Mihaita, A.S. and Cai, 2019. On the other hand, simulation-based methods are commonly used in the literature to estimate PT patronage patterns. These methods involve creating simulations following assignment models that replicate real-world scenarios, allowing researchers to observe and analyze the change in behaviour and dynamics of PT systems (see Zhao et al., 2022; Shafiei et al., 2021b; Wen et al., 2018; Saberi et al., 2018; Zhao, 2019; Shafiei et al., 2020) so that the essential regarding the impact on the network can be captured and modelled. Some propose as well optimisation techniques under disruptions to ease impact (see Mao, Mihaita, and Cai, 2019).

In this research chapter, we analyze the PT patronage patterns using real smart-card data and compare the normal versus incident circumstances in order to estimate the impact on PT users. As an observation, the PT patronage appears to have a daily repetitive occurrence, and we draw inspiration from signal analysis principles. Therefore, we employ a Fourier transform function to fit the data and filter out the noise, enabling us to model the underlying pattern of PT patronage effectively. In the past, one study utilized the method to model road traffic patterns by analyzing traffic volumes derived from mobile data (see Abera and Hailemariam, 2018). Another study employed the graph Fourier transform (GFT) for a similar purpose (see Chindanur and Sure, 2018). However, these studies primarily focused on modelling private vehicle traffic and not public transport patronage. By establishing a robust model capable of accurately representing traffic patterns, we take an additional step in our research by not only modelling the PT patronage pattern but also applying the model to identify the effects of incidents on patronage.

4.1.2 Chapter Contributions

To summarise, the main theoretical and methodological contributions of this chapter are the following:

- We introduce a novel method for dynamically modelling PT patronage patterns by incorporating the Fourier transformation to effectively reduce the influence of noise and enhance the accuracy and reliability of the model;
- We employ the frequency domain following the Fourier transform to segregate the components of PT patronage patterns, enabling us to identify and isolate the significant elements within these patterns;
- We apply a synergistic approach that combines analytical techniques with data-driven methods to identify the impact of incidents on PT passengers;
- We perform the ability of multiple measuring metrics, such as correlation measures (Pearson's correlation coefficient), distance measures (Chebyshev distance, Wasserstein metric, Minkowski difference and Cosine similarity) and statistical tests (change, Percentage change and Symmetric percentage change);
- We propose a new application by integrating big data resources, among which GTFS data, smart-card and incident log data when excavating the information for analysing the network vulnerability.

This chapter is organised as follows. In Section 4.2, the dynamic PT patronage pattern model considering the Fourier transform is discussed, and the details of incident impact identification metrics are also included. The application of the proposed methods to a real network is presented in Section 4.3, and the results of the case study are demonstrated in Section 4.4, where detailed modelling processes are demonstrated. Finally, the research conclusion and the future directions are provided in Section 4.5.

4.2 Methodology

4.2.1 Entity of PT patronage

Number of boarding and alighting

Given smart-card data, we can produce the number of boarding and alighting for each PT stop at each time spot or for each time interval, in order to represent the PT patronage. Each record r in the smart-card data, based on the availability for use in this research, can be expressed as:

$$r(i, u_i, j, u_j, b, d), \tag{4.1}$$

where each parameter in a record r represents the tap-on stop i, tap-on time u_i, tap-off stop j, tap-off time u_j, the bus number b and the date of the recording d. To obtain the number of boarding and alighting for a certain stop $i, i \in \{1 \dots I\}$ during a period of time (τ), we only need to count the number of records, denoted as $N_i^{boarding}(\tau_a), a \in \{1 \dots A\}$, where a represents the a^{th} time interval of a day, A is the total number of time interval defined for a day; similarly, the number of alighting at a specific stop during a time interval can be expressed as $N_j^{alighting}(\tau_a)), j \in \{1 \ldots I\}$. The number of boarding and alighting people under an impact of disruption can be expressed as $N_i^{boarding'}(\tau_a)$ and $N_j^{alighting'}(\tau_a))$. The total number of boarding persons, counted during the time interval τ can be denoted, according to Iverson bracket notation, as:

$$N_i^{boarding}(\tau_a) = \sum_{n=1}^{N} [u_i(n) \in \tau_a], \tag{4.2}$$

where $u_i(n)$ represent the n^{th} tap-on time recorded in the data set; if this time belongs to τ_a is true, then $[u_i(n) \in \tau_a]$ is 1; otherwise, this record is not counted.

Based on the given definition, we observe that the patronage data relying on boarding and alighting focuses on the bus stops where passengers get on and off while ignoring the passed stops along a trip. However, when it comes to trips passing through an area affected by an incident with persons already boarded at other stations, this counting method fails to account for the actual number of passengers impacted. Therefore, to accurately determine the number of disrupted passengers specifically caused by the incident, an in-vehicle passenger count is also necessary. This method allows us to count the number of passengers affected within the impacted area and provides a more precise measure of the disruption's impact.

Number of in-vehicle passengers

Since we hold the information around the time and location when passengers got on/off the bus, we can define the number of in-vehicle passengers for each time interval at each PT stop. For each smart-card data record r, we are able to calculate the number of time intervals (τ) through which this trip passes from the start until the end:

$$K = \frac{u_j - u_i}{\tau}, k \in 0 \ldots K, \tag{4.3}$$

For each τ passed by a trip, a passenger is counted at each $k^t h$ interval τ_k. K is estimated by rounding the number of τ in order to improve processing accuracy. In our study, we consider 15 minutes as the time interval, so if a record of patronage starts at 8 am and ends at 9 am, instead of counting this record by boarding time (interval 8:00-8:15) once, we count it four times, for 8 am, 8:15, 8:30 and 8:45 slot to represent that this passenger is inside the vehicle from 8 am to 9 am.

According to the above definition, each record in the smart-card data set has added another element regards to the in-vehicle time, represented by $v_{i,k}$, where k denotes the k^{th} time interval τ that this trip is passing through:

$$v_{i,k} = u_i + k\tau. \tag{4.4}$$

Therefore, each record in the data set is updated by adding an in-vehicle time $v_{i,k}$ which equals the u_i, and $K - 1$ new records are filled in the data set which corresponding to their boarding location i, boarding time u_i, in-vehicle time $v_{i,k}$, PT number b and date d, denoted as:

$$r(i, u_i, v_{i,k}, b, d). \tag{4.5}$$

Following the definitions, the number of in-vehicle passengers becomes:

$$N_i^{in-veh}(\tau_a) = \sum_{n=1}^{N} [v_{i,k}(n) \in \tau_a]. \tag{4.6}$$

4.2.2 Modelling PT patronage via the Fourier Transform

By taking into account the total number of passengers onboard, it now becomes possible to identify the number of affected passengers by comparing the change between a travel pattern on the day of the incident and on a typical day. To estimate the travel pattern, we require to use the incident log data, including information regarding incident time, duration, and location. In order to obtain the travel pattern for a typical day, one can apply either the traditional approach of averaging the patronage counts during non-incident days, or the Fourier transform and filter out noises from daily travel patterns.

In this work, we propose to use the Fourier Transform to analyse time-dependent signals in the frequency domain; it is a tool for decomposing a complex and repetitive behaviour pattern by summing up the sines and cosines functions. This concept can be adapted for the PT patronage estimation because such a patronage pattern is time-varying and repetitive over a certain time, exhibited by distinct peaks during the morning and afternoon periods. The seasonality enables the patronage patterns to be predictable by using the discrete Fourier transform function. According to Brunton and Kutz, 2021, the frequency-domain function following the discrete Fourier transform can be expressed as:

$$f(\tau) = \frac{\alpha_0}{2} + \sum_{h=1}^{H} (A_h \sin(h\omega\tau + \varphi_h) + B_h \cos(h\omega\tau + \varphi_h)). \tag{4.7}$$

where A_h is given to describe the amplitude of the *sine* function while B_h is the amplitude of the *cosine* function. These two parameters indicate how much sine and cosine functions should be included to estimate the function of the travel pattern. $\omega\tau$ indicates the frequency component of the trigonometrical function and φ_h is the phase of the trigonometrical function.

By decomposing the patronage pattern function into multiple sine and cosine waves, it becomes possible to convert the data from the time domain to the frequency domain and following this, capture the regularity exhibited on the frequency scale. In the frequency domain, we have the magnitude spectrum by frequency (or the power spectrum in signal analysis). Those frequencies with significant magnitudes are considered major frequencies, which indicate the dominant power contained within the signal, while frequencies with low magnitudes are treated as noise. Utilizing this information, we can filter the useful data from noise data based on magnitude. This is how we de-noise or approximate any arbitrary function by summing up a determined set of trigonometric functions.

4.2.3 Measurements of impact

There are several methods for measuring the impact of the incident according to the change of patronage with and without the incident. The options for measurement include correlation measuring, such as Pearson's correlation coefficient; the metrics related to distance, such as the Chebyshev distance, the Wasserstein metric, the Minkowski difference and Cosine similarity, as well as the statistical tests, such as the change, the Percentage change and the Symmetric percentage change.

1) Change:

$$I^{change}(N, N') = N - N', \tag{4.8}$$

where N represents a set of patronage counts (number of in-vehicle passengers, as described in Equation 4.6) for a typical day and N' for the incident day.

2) Percentage change and symmetric percentage change: To avoid the problems triggered by the value for the disrupted situation being zero, we adopt the symmetric percentage change, as well, which is given as Equation 4.10.

$$I^{Pchange}(N, N') = \frac{N - N'}{N + \lambda} \times 100\%, \tag{4.9}$$

where λ is a smoothing factor used to avoid computing problems when dividing by zero.

$$I^{Pchange}(N, N') = \frac{N - N'}{\frac{N+N'}{2}} \times 100\%, \tag{4.10}$$

By using the symmetric percentage change, the result that approaches either 2 or -2 means that there is no similarity between the two data sets, while if the result ranges to zero, it indicates that these two data sets have high similarity.

3) Cosine similarity: which is the dot product of the number of in-vehicle passengers affected by an incident N' and the number of counts for a typical day, as:

$$I^{\text{Cosine}}(N, N') = \frac{N \cdot N'}{\|N\|\|N'\|} = \frac{\sum_{m=1}^{m} N_m N'_M}{\sqrt{\sum_{m=1}^{M} {N_m}^2}\sqrt{\sum_{m=1}^{M} {N'_m}^2}}, \tag{4.11}$$

4) Chebyshev distance: This is defined as the maximum distance along any coordinate dimension which measures the greatest discrepancy in values between the corresponding coordinates of the vectors being compared:

$$I^{\text{Chebyshev}}(N, N') = \max_m(|N_m - N'_m|), \tag{4.12}$$

5) Wasserstein distance: The metric serves as a distance function defined between probability distributions on a given metric space. N and N' are two measures on a metric space $\mathbb{R} \times \mathbb{R}$; the Wasserstein distance between these two measures is defined as the integration of the distance between any two matched points times the amount of the mass of moving from one point to another. Thus Wasserstein distance is given by:

$$I^{\text{Wasserstein}}(N, N') = inf_{\pi \in \Gamma(N,N')} \int_{\mathbb{R}\times\mathbb{R}} |N - N'| d\pi(N, N'), \tag{4.13}$$

where π is the joint probability measure on $\mathbb{R} \times \mathbb{R}$ with marginals N and N'.

6) Minkowski difference:

$$I^{\text{Minkowski}}(N, N') = \|N - N'\|_p == (\sum |N - N'|^p)^{1/p}, \tag{4.14}$$

where p is the order of the norm of the difference between N and N'. When $p = 1$, the Minkowski distance is the same as the Manhattan distance, while $p = 2$, such distance is the same as the Euclidean distance. The value of p is 3 in this chapter based on the experiment's comparison result.

7) Pearsons correlation coefficient (PCC): PCC is a way of quantitatively measuring the linear correlation; it assesses the extent to which changes in one variable are associated with corresponding changes in another variable, both in terms of direction and magnitude.

$$I^{\text{PCC}}(N, N') = \frac{\sum(n_m - \overline{n})((n'_m - \overline{n'})}{\sqrt{\sum((n_m - \overline{n})^2 \sum(n'_m - \overline{n'})^2}}. \tag{4.15}$$

4.3 Case study

4.3.1 Network characteristics

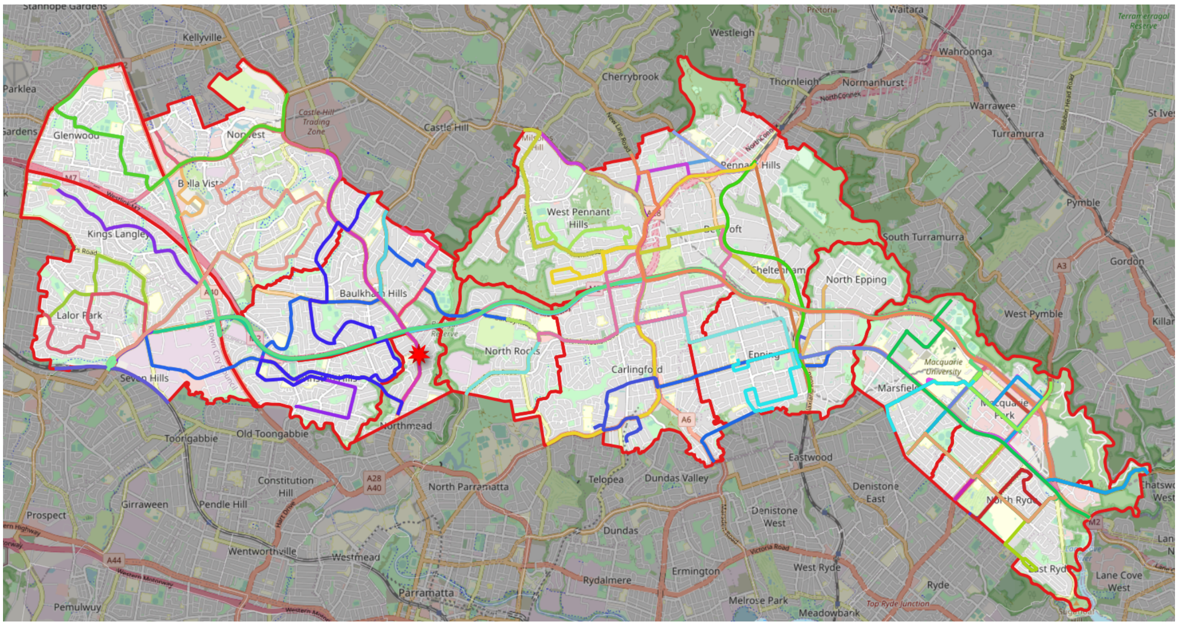


FIGURE 4.1: Map of the Sydney M2 area showing road and PT networks.

The study focuses on the zones in the North-West region of Sydney, encompassing the M2 motorway and comprising significant residential and commercial sectors. The geographical extent of this area, depicted in Figure 4.1, aligns with the boundaries defined by the Statistical Area Level 2 (SA2) Australian Bureau of Statistics, 2021 in digital mapping, where there are 79 bus lines consisting of 3,799 bus stops.

4.3.2 Sample of a hypothesised incident

To evaluate the feasibility of metrics on identifying the impact of patronage patterns due to the traffic disruption, we consider comparing the modelled typical profile with a hypothesised incident one and reflecting the impact by metrics as mentioned in Section 4.2.3.

The hypothesised incident scenario is created following the details of a real incident (as introduced in Section 4.3.3 below). However, rather than directly comparing the day of the incident with a modelled typical day, we measured the difference in passenger count during the incident duration. We then incorporated this change into the patronage of the modelled typical day, effectively creating a hypothesis day of the incident. The details of the disruption are designed to mirror the scenario presented in Section 4.3.3, with a start time of 2017-04-05 at 10:03:59 and a duration of 50 minutes. To simplify the modelling process, we approximate the time unit to 15 minutes. This makes the hypothesised incident start from 10:00:00 (time interval 40) until 10:45 (time interval 43). During the disruption, the impact manifests as an increase in patronage, with 2,397 additional passengers distributed across three 15-minute intervals (927, 790 and 680 passengers, respectively). To assess the impact, we manually adjust the count within the modelled typical day scenario for each time index interval. This allows us to observe the resulting changes reflected in the impact measuring metrics.

4.3.3 Sample of a real incident

In order to measure the impact of an incident, the sample incidents selected from the incident log data set should follow the considerations: such incident duration is long enough (at least 30 minutes) to be able to display the impacts through the PT patronage; such incident is away from the PT-only lane; because the impact on patronage for an isolated PT could be minor Zhao, 2019; such incidents should be located in prominent residential and commercial sectors considering the uneven distribution of patronage data. A sufficiently large dataset of patronage is necessary to ensure a discernible impact; such an incident possesses a substantial potential to impact PT patronage significantly.

In this research, we have selected and tested multiple sample incidents to ensure that our findings are applicable to a wide range of scenarios. However, for the purpose of showcasing the results and the limited space, we specifically selected this particular incident in this chapter:

- Start time: 2017-04-05 10:03:59
- Duration: 50 min
- Type: Bus Breakdown

More data analysis results can be found in supplementary material Online-supplement, 2022. All results displayed in the following sections correspond to this sample incident.

4.4 Results and discussion

4.4.1 Profiling a typical day patronage pattern using the real data

To capture the count pattern for a typical day, we explore two methods. The first method involves calculating the average of multiple non-incident days, where the day of the week matches that of the incident. This approach takes into account the observation that different days of the week exhibit distinct patronage patterns, as shown in Figure 4.3. In the case of an incident occurring on Thursday, 5 April 2017, we gather the remaining non-incident Thursdays in April 2017 and compute the average count for each day. As mentioned in Section 4.2.1, to optimize data input and streamline data processing, we have chosen to calculate the total number of in-vehicle passengers at each PT stop instead of separately considering boarding and alighting counts.

This approach effectively doubles the data size by incorporating the combined number of tap-on and tap-off events, as depicted in the three plots in Figure 4.2. This consolidation simplifies the analysis while maximizing the available data as follows: *a*) Figure 4.2 demonstrates the number of in-vehicle passengers with and without an incident, where the red dash line in the figure highlights the duration of the sample incident. Comparing this figure with *b*) and *c*) in Figure 4.2, which is generated by using the number of boarding and alighting, we observe that the total number of in-vehicle count on a typical day is 17,779 and that on the incident day is 18,014, whereas the sum of count for tap-on on a typical day is 9,310, and on the incident day is 9,565; additionally, the tap-off on a typical day is 11,224, and the number of tap-off on the incident day is 11,342. This comparison highlights that if we would utilize only the original tap-on data then we would have approximately 47% of trips being ignored when compared to using the in-vehicle passenger count. Similarly, when using the original tap-off data, approximately 37% of trips are overlooked.

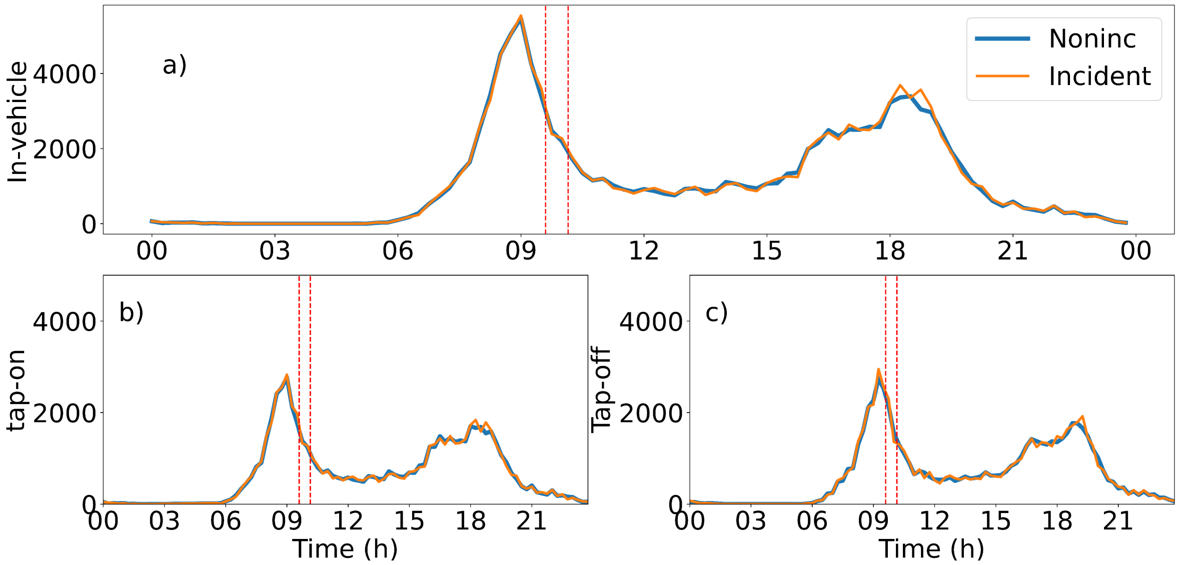


FIGURE 4.2: Patronage on a typical (non-incident) versus incident day by a) in-vehicle passengers, b) tap-on passengers, c) tap-off passengers.

4.4.2 Modelling a typical day patronage pattern using the Fourier transform

All sub-figures in Figure 4.2 present challenges in distinguishing between a typical profile and an incident profile. Despite observing numerous fluctuations and variations between the two profiles in both figures, it remains challenging to discern which changes are specifically attributed to the incident under investigation. This difficulty comes from the complex nature of the system, making it hard to ascertain the impact of the incident on the PT patronage pattern. The complexity of the system necessitates the purification of patterns by removing unnecessary noise. This requirement motivates the application of the Fourier transform, as it allows for the decomposition of complex patterns into individual simple and determined components. By examining the performance of each component separately, we can approximate the incident's impact as noise within the seasonal travel pattern.

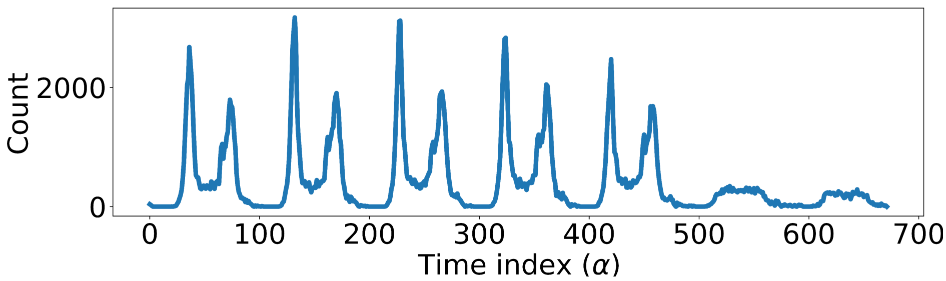


FIGURE 4.3: Weekly time domain depicted using real smart-card data.

For the specific incident under consideration, we collect a week of data encompassing the incident day and apply the Fourier transform to convert the time domain into the frequency domain. After removing the noise components in the frequency domain, we reverse the denoised frequency domain back to the time domain to form the profile for a typical day. From this denoised time domain, we select the travel pattern for Thursday as our representative pattern. According to the weekly travel pattern depicted in Figure 4.3, we only select the data for Monday to Friday. These weekdays show a similar pattern, which makes them suitable for our analytical purposes. To better visualise the tendency in the plot, we convert the date and time information into date-time-index (α) by 15 minutes.

Frequency domain: After applying the analytical process outlined in Section 4.2.2, we convert the seasonal time domain into the frequency domain. The frequency domain representation is depicted in Figure 4.4, with the left plot showcasing the overall magnitude spectrum and the right plot specifically highlighting the dominant components. In this representation, a frequency of 1/15 minutes is utilized, implying that each frequency value corresponds to the number of occurrences of a repeating event within a 15-minute interval. The properties for the top three components (highlighted by the red dots in the right sub-plot of Figure 4.4) are displayed in Table 4.1 and processed by adding period information to make the data more understandable.

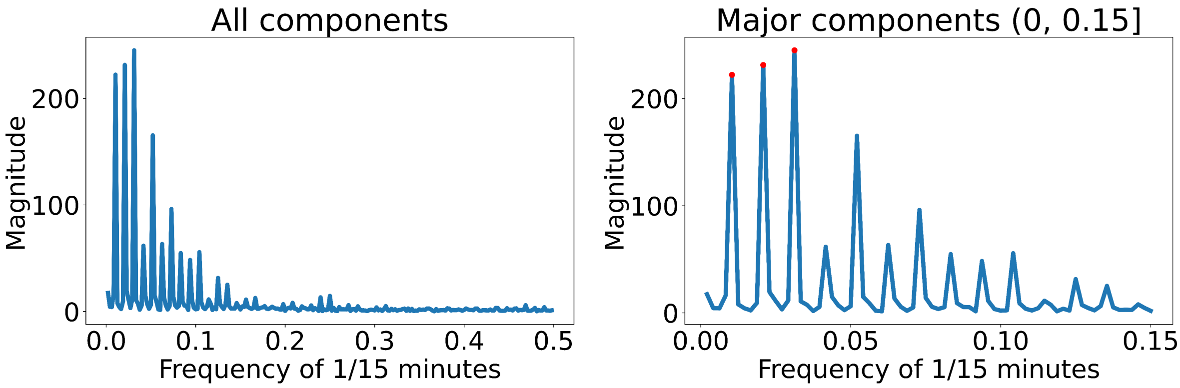


FIGURE 4.4: Frequency domain representation.

Period and seasonality: To enhance comprehension, we convert the frequencies to periods (see Table 4.1) by taking the reciprocal of the frequency, represented in hours and days. Each period value means the duration of time of one cycle in a repeating event. This conversion facilitates a clearer understanding of the data. The highest magnitude (given as the first row of data shown in Table 4.1

TABLE 4.1: Samples of the Fourier transform outcomes.

Notes: FFT result is the Fourier transform outcome, in complex value; Mag is magnitude; Freq is frequency per 15 minutes; Pd(h) is period by hour; Pd(d) is period by day.

No.	FFT result	Mag	Freq	Pd (h)	Pd (d)
1	53423-104789j	245	0.031	8	0.33
2	-70066+86141j	231	0.021	12	0.50
3	-98011+41997j	222	0.010	24	1.00

matching the highest bar in the right figure of Figure 4.4), representing the most notable seasonal patterns, exhibits a period of 8 hours (in Table 4.1), indicating that this event repeats every 8 hours, aligning with the off-peak hours. The second highest magnitude corresponds to a period of 12 hours, reflecting the morning and afternoon peaks that repeat every half day. The event with the third highest magnitude repeats daily (repeated every 24 hours), suggesting it captures the overarching function that describes the seasonality of the patronage pattern throughout the day.

Decomposition of travel pattern: The Fourier transform, as defined in Section 4.2.2, involves transforming a function into a series of increasing high-frequency periodic functions. In the context of PT travel patterns, we can decompose the pattern into repetitive sub-functions. By examining the periods obtained through the Fourier transform, as shown in Table 4.1, we can match the period information to the actual seasonality. At this stage, we can effectively filter out the noise and capture the dominant characteristics of the travel pattern. In other words, the Fourier transform allows us to isolate and analyze the significant ingredients of the PT travel pattern while disregarding irrelevant fluctuations or noise. Figure 4.5 demonstrates the major frequency components in the PT patronage pattern plotted by sines and cosines. We can observe that the periodic peaks of each sub-function align

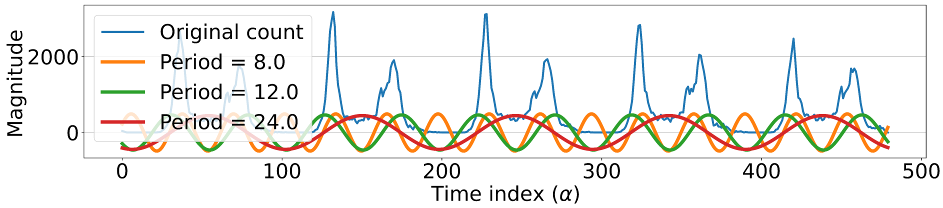


FIGURE 4.5: Top three frequency components in weekly PT patronage pattern.

with the count pattern depicted by the actual count data (represented by the blue wave in Figure 4.5). The orange wave, with a period of 8 hours, corresponds to three off-peak periods in the time domain representation (blue wave). The green wave, with a period of 12 hours, matches two peak hour periods. Lastly, the red wave, with a period of 24 hours, captures the day and night periodicity. This alignment demonstrates how the sub-functions derived from the Fourier transform effectively capture the characteristic patterns present in the actual count data.

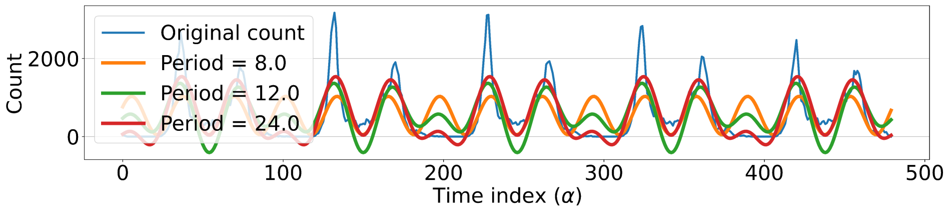


FIGURE 4.6: Time domain by reverse Fourier transform considers top three components.

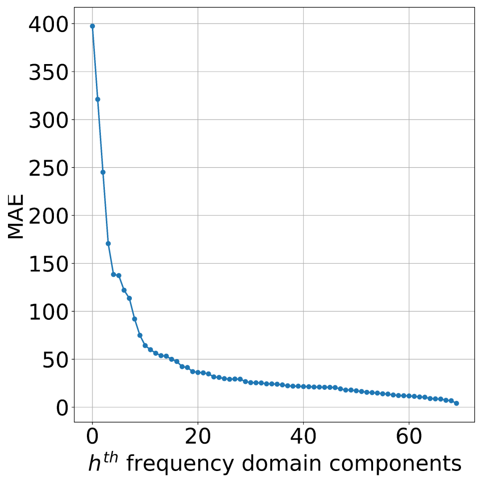


FIGURE 4.7: Mean absolute error (MAE) of various extent of filtered time domain by reverse Fourier transform and real data.

Given that the components derived from the Fourier expansion exhibit harmonic frequencies, phases, and amplitudes, we can cumulatively combine them to construct the desired approximate function, as shown in Figure 4.6. In the time-domain representation, the yellow line represents the pattern considering only the top component (No.1 in Table 4.1); incorporating the top two components results in the result plotted in green, while considering all three top components yields the red line in the time-domain representation.

As we incorporate additional components and transform the frequency domain back to the time domain using the inverse Fourier transform, the modelled pattern increasingly aligns with the original pattern. This is proved by results in Figure 4.7.

Filtered time domain representations: Upon transforming the time-domain representation into a frequency-domain one, we have broken down the travel pattern into several periodical functions. Our next step is to determine which periodic functions should be incorporated to construct a modelled travel pattern that is not only clean in structure but also accurately reflects the tendency of patronage pattern. Therefore, we conduct a sensitivity test to evaluate the impact of different periodic functions on the model performance. The outcome of the sensitivity test is presented in Figure 4.7, which extends the findings presented in Figure 4.6. This figure illustrates the modelling performance ranging from incorporating the top component to including all 70 components.

In Figure 4.7, we can observe that adding more component functions during the modelling process leads to a reduction in mean absolute error (MAE), which indicates an improved model performance. Notably, the curve shows a deeper slope within the range of 0 to 10 compared to that between 20 and 30 and further. This implies that the advantages gained by including less than 10 components are insufficient while including more than 30 components appears to be unnecessary. So the ideal range should be between 10 to 20.

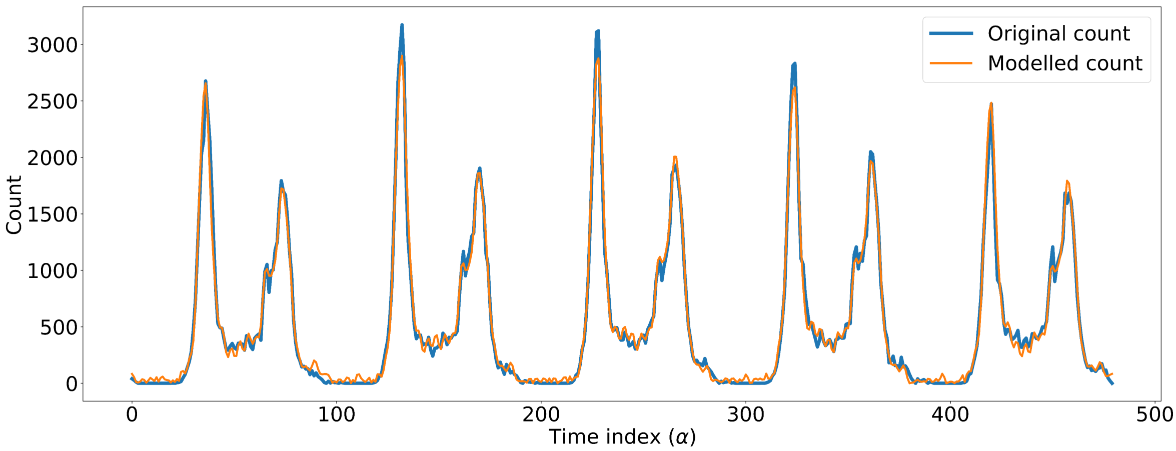


FIGURE 4.8: Time domain by reverse Fourier transform considers top 15 components.

Consequently, at this stage, we consider the repetitive components and their meaning in the real world to determine whether the frequency/period should be included, as the explanation for Table 4.1. And we decided to model the PT patronage pattern by using the top 15 components while excluding the remaining components. To showcase the efficacy of our modelling approach, we employ the inverse Fourier transform to generate a time-domain representation of the model, as depicted in Figure 4.8.

Nevertheless, given that the primary objective of this chapter is to assess the impacts of disruptions on the PT patronage pattern, we must consider if the modelled typical profile can reflect the impacts effectively. To this end, we enhance the sensitivity test by including a similarity test between the modelled typical and incident profiles; the detailed explanation can be found in Section 4.4.3.

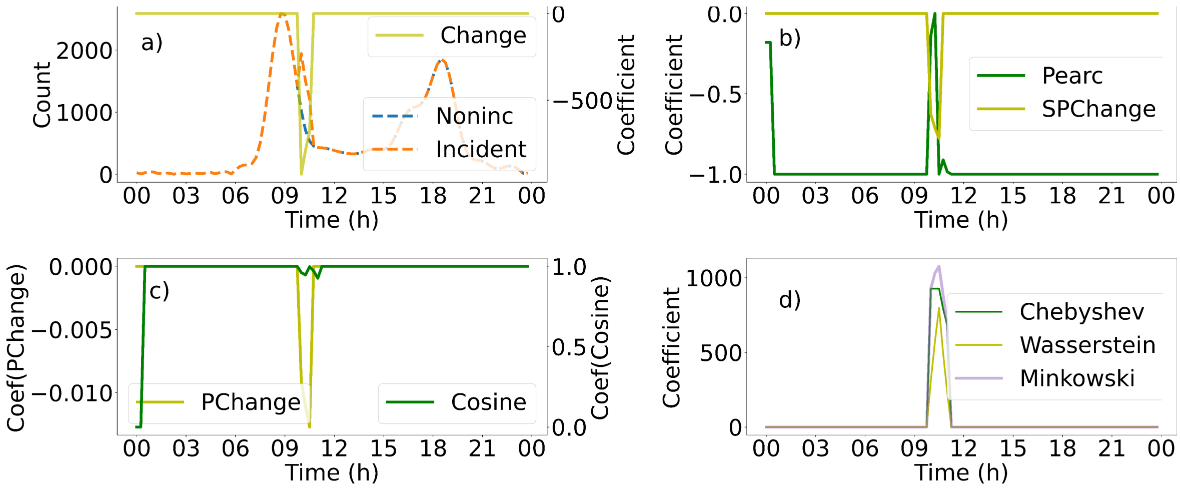


FIGURE 4.9: Performance of metrics considers a hypothesised incident profile.

4.4.3 Incident impact measurement

Measure the impact of the hypothesised incident

Through the application of a hypothesized incident, we are able to control the level of noise (variables), enabling us to assess the performance of the metric directly.

The presented Figure 4.9 illustrates the outcomes of all metrics (as shown in subsection 4.2.3) employed to evaluate the resemblance between the profile of the modelled typical day (using 15^{th} significant periodic components) and the incident day. In this figure, we can observe that all metrics perform well when identifying the change in patronage.

Measure impact using real incidents

However, when measuring the change using data with noises (measuring the impact using real incident and modelled typical profile), the count change Figure 4.10-a) and symmetric percentage change Figure 4.10-b) follow the flow of the peak and off-peak hours; the percentage change, Cosine similarity Figure 4.10-c) and Pearson's correlation Figure 4.10-b) prove to be less effective as they closely align with the large percentage change. In contrast, distance measurements in Figure 4.10-d) exhibit better performance. Out of the Chebyshev distance, Wasserstein distance, and Minkowski difference metrics, the Chebyshev distance yields a more straightforward shape that effectively illustrates the observed trend. Consequently, we solely present the results utilizing the Chebyshev distance in the subsequent visual representations.

As mentioned in the last paragraph of Section 4.4.2, it is crucial to assess the ability of the modelled profile to capture the impact accurately. This approach enables us to determine the optimal combination of components that can effectively represent the PT patronage pattern for a typical day so that we are able to encapsulate the influence of a real-life incident effectively.

The result is displayed in Figure 4.11 following metrics of Chebyshev distance. It is apparent that incorporating the top three components yields the most satisfactory outcome. However, upon visualising the measurement results over time, as depicted in Figure 4.12-a), we notice substantial disparities during peak hours in the morning and afternoon. Thus, solely accounting for the similarity between

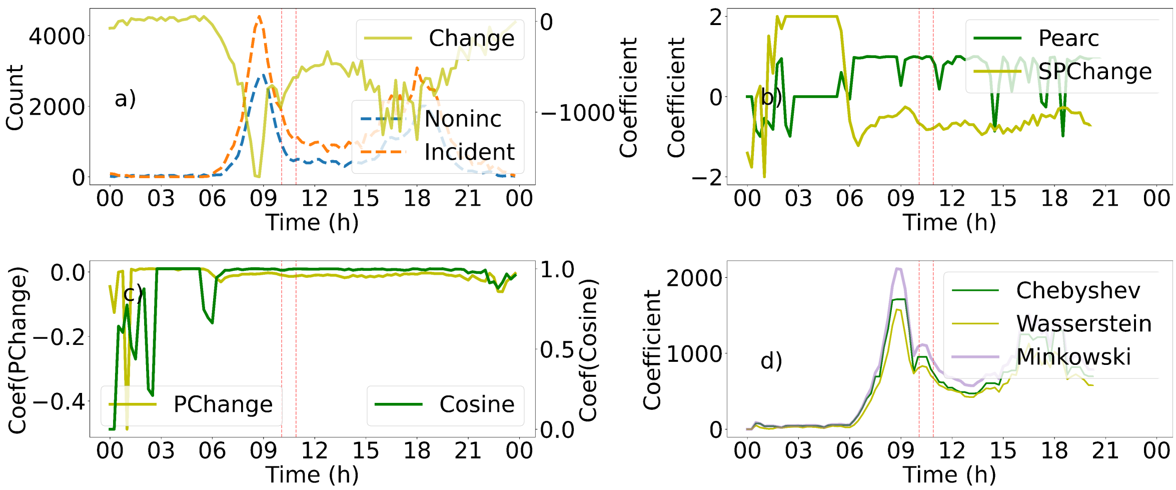


FIGURE 4.10: Performance of metrics considers the real incident profile.

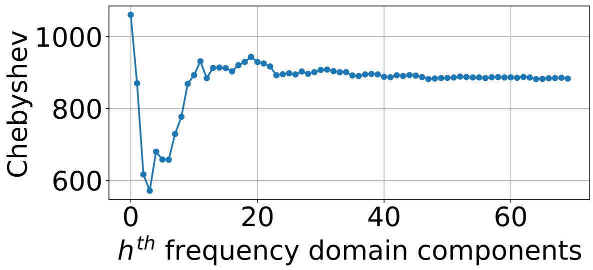


FIGURE 4.11: Modelling sensitivity test

an ordinary day and an incident day might not be adequate to best encapsulate the disruption's effect on the PT patronage pattern; this implies that the model generated using the top three components may not necessarily be the most optimal choice for impact identification purposes.

Figure 4.12 displays a series of snapshots that show the performance of identifying the disruption impacts. Each snapshot indicates the performance of adding a different number of frequency domain components when modelling; the model performance is reflected by the Chebyshev distance (green line). For example, Figure 4.12-a) shows the result when adding the top three components, where the similarity of a typical (blue line) and incident (orange line) day is evaluated by the Chebyshev distance. The red dash lines highlighted the incident duration. However, it is difficult to discern any noticeable

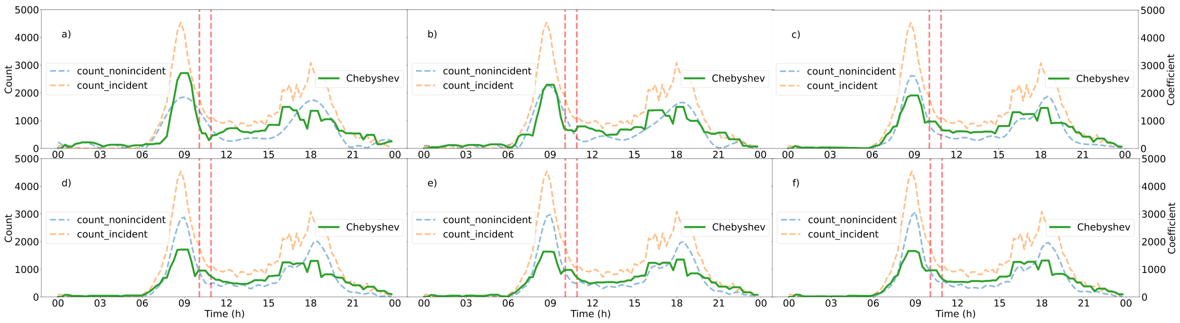


FIGURE 4.12: Performance of real incident impact detection by including the n^{th} frequency domain components: a) 3^{th}, b) 6^{th}, c) 11^{th}, d) 15^{th}, e) 18^{th}, f) 30^{th}.

differences during the incident in this sub-figure. Instead, the peak measurements are predominantly evident during the morning and afternoon peak hours.

When we include a greater number of components in the modelling of the typical day, for instance, as shown in Figure 4.12-c), where the top 11 components are considered, we can observe a distinct sub-peak during the incident duration (as indicated by the two red dashed lines). However, it's important to note that the peak measurements persist during the morning and afternoon peak hours, which can be attributed to the significant flow and substantial fluctuations that typically occur during these times.

The observed performance trend in Figure 4.12 aligns with the findings presented in Figure 4.11. When incorporating the top three components, the model primarily captures the characteristics of the morning and afternoon peaks. However, by including the top 11 components, the model is able to capture not only the impact during peak hours but also the concurrent noise resulting from the incident. Furthermore, as more components are added, the model's performance remains relatively stable.

4.5 Conclusion

The proposed method in this chapter aims to dynamically model the PT patronage patterns and identify the impacts of road incidents on PT users. The proposed method applies the Fourier transform to decompose complex patterns into distinct waves; this allows the dominant components of the patronage pattern to be capturable and used as the reference (typical) profile for traffic analysis. One specific application showcased in this chapter is impact identification. The presence of peak hours makes it challenging to capture the current incident impacts accurately. However, through an enhanced sensitivity test that considers the performance of impact identification, we can improve the modelling ability of the typical day. This improvement enables us to capture the current impact effectively. Multiple sample incidents are tested using this method, and the results obtained are robust. However, due to word limitations, only the results of one sample incident are presented in this chapter. More data analysis results can be found in supplementary material Online-supplement, 2022.

The proposed modelling method allows us to identify the optimal typical profile. However, the current model's effectiveness relies on a substantial amount of count data and certain disruptions (such as prolonged durations and locations far from the PT-only lane) to generate precise results. These limitations present opportunities for enhancement in future research. While zonal data is utilized in the analytical processes, we believe that focusing solely on nearby stop data, for example, for analysis could significantly enhance the model's ability to identify and assess the impact using this particular modeling approach. As for future directions, further exploration can focus on decomposing the patterns and effectively identifying the noise generated by disruptions, such as recurring congestion or incidents. This would allow for quantifying the impacts of these disruptions by using wave functions, for example. Additionally, the model has the potential to be expanded to incorporate spatial analysis. By introducing an additional spatial dimension, it becomes possible to capture the evolution of impacts based on location. This extension would provide valuable insights into the spatial dynamics of disruptions and their effects on PT patronage patterns.

Chapter 5

Traffic disruption modelling with mode shift in multi-modal networks

A multi-modal transport system is acknowledged to have robust failure tolerance and can effectively relieve urban congestion issues. However, estimating the impact of disruptions across multi-transport modes is a challenging problem due to a disaggregated modelling approach applied to only individual modes at a time. To fill this gap, this chapter proposes a new integrated modelling framework for a multi-modal traffic state estimation and evaluation of the disruption impact across all modes under various traffic conditions. First, we propose an iterative trip assignment model to elucidate the association between travel demand and travel behaviour, including a multi-modal origin-to-destination estimation for private and public transport. Secondly, we provide a practical multi-modal travel demand re-adjustment that takes the mode shift of the affected travellers into consideration. The pros and cons of the mode shift strategy are showcased via several scenario-based transport simulating experiments. The results show that a well-balanced mode shift with flexible routing and early announcements of detours so that travellers can plan ahead can significantly benefit all travellers by a delay time reduction of 46%, while a stable route assignment maintains a higher average traffic flow and the inactive mode-route choice help relief density under the traffic disruptions.

This chapter is based on an edited edition of the following article: Zhao D, Mihaita AS, Ou Y, Shafiei S, Grzybowska H, Qin K, Tan G, Li M. Traffic disruption modelling with mode shift in multi-modal networks. IEEE ITSC 2022, Macao, China.

Contents

5.1 Introduction

5.1.1 Background and motivation

Resilient cities have recently embraced a fully-connected multi-modal transport network that gives travellers more freedom when choosing when, where and how to travel. However, multi-modal urban environments are also vulnerable due to the lack of tolerance against an ever-growing population, an expanding travel demand, a high private car ownership, deficient transport design, inadequate traffic control and flawed travelling or driving behaviour (see Rahman et al., 2021).

To improve the efficiency of the transport system at a large scale, the encouragement of a travel behaviour change and active mode shift is an encouraging option studied recently (see Ettema et al., 2016). Many other research studies reinforce this initiative by providing substantial evidence via data-driven, or simulation-based approaches (see Wen et al., 2018; Mihăiţă, Dupont, and Camargo, 2018; Mao et al., 2021). The data-driven approaches capture the real traffic behaviour before and after disruptions, and some applications are used in programs such as: INPHORMM, TAPESTRY or Travel Smart (see Ma, Mulley, and Liu, 2017). Other early studies revealed the value of public transport by investigating the change of traffic states (e.g. section flows, traffic volumes or travel times) and proposed an entire public transport service removal when massive public transport disruptions occur or when service is suspended (see Moylan, Foti, and Skabardonis, 2016; Adler and Ommeren, 2016). Few studies that consider a simulation approach mention that the change in the level of congestion before and after the removal of public transport services would clarify the significance of public transport (see Nguyen-Phuoc et al., 2018). More recently, the unprecedented COVID-19 pandemic has heavily modified the travel demand and provided evidence with regards to the impact of traffic demand across all mode shifts in a city (see Das et al., 2021).

Challenges: All previous studies solve the mode choice problem before departing, and most publications provide modelling methods from a macroscopic or a mesoscopic level based on a statistical analysis. There is little research into investigating the benefits of an active mode shift from a dynamic microscopic perspective and its impact when traffic disruptions occur. A significant gap is present due to the lack of data regarding the impacted demand under incidents and active mode shifts. Some studies rely on surveys or a stated preference obtained ahead of trips to obtain the number of impacted travellers or the number of mode and route shifts (see Auld et al., 2020; Guzman et al., 2021). However, we emphasise identifying the impacted origin-to-destination (OD) trips affected by disruptions in a simulating model, and the change of mode and route choice that leads to a demand change is employed for evaluating the impact on network performance in our work.

Apart from the lack of data, quantifying the impact of disruptions is also a major challenge; some research studies have analysed the change of trip-based mean delay, mean speed (see Aftabuzzaman, Currie, and Sarvi, 2010) or travel time (see Nguyen-Phuoc et al., 2018). However, such indicators can hardly differentiate the impact from the general traffic (e.g. recurrent congestion) or from traffic control strategies. To address this issue, we work across several indicators versus baseline conditions in order to evaluate the efficiency of the proposed ones.

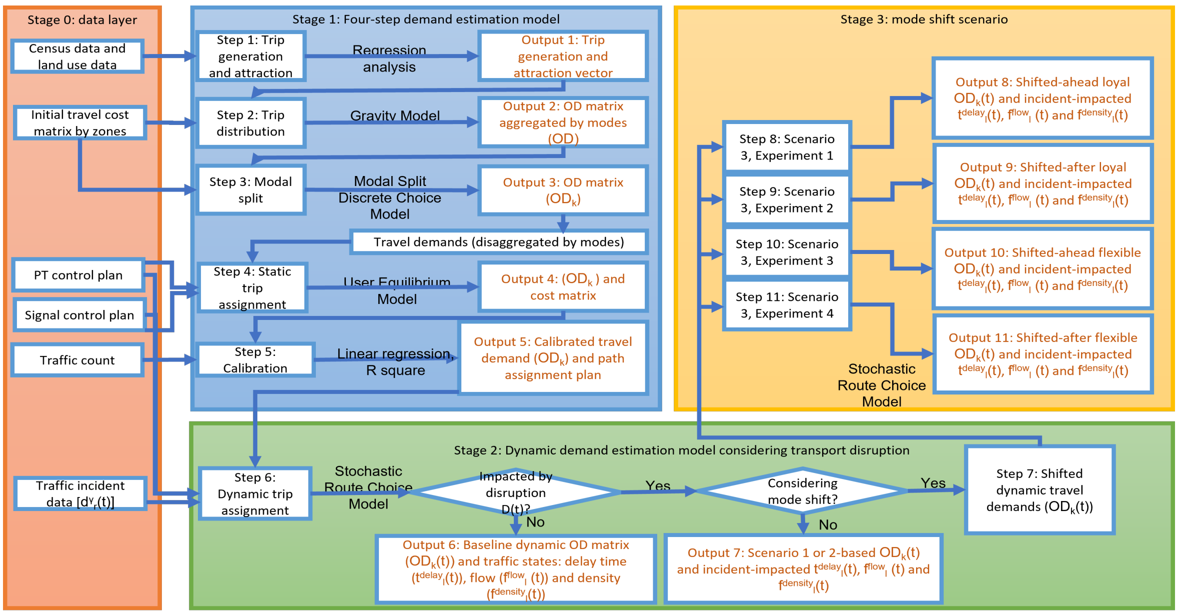


FIGURE 5.1: Framework of our proposed multi-modal transport network modelling under disruptions.

Another major challenge of dynamically simulating the mode shift is the lack of dynamic demand data and the method of integrating the OD estimation across different transport modes in order to identify the impacted trips. Most previous research studies only consider a single-mode (see Thompson et al., 2019), while some research studies model car-based transportation versus public transportation differently (see Hussain, Bhaskar, and Chung, 2021). Extensive evidence considers the OD estimation from total generation and attraction data based on the gravity model. This method has been largely developed with the improvement of mathematical, analytical and computational skills. However, the large potential of the gravity model approach in the transport field has not yet been fully explored, as most research studies attempted to investigate the OD matrix for a single transport mode, mostly cars. There is still a need to consider the influence of other transport modes when mode splitting and trip assigning under a multi-modal public transport environment. The challenge of integrating the OD matrices of various public transport modes with that of private vehicles is still unsolved.

5.1.2 Chapter Contributions

In this chapter, we propose an integrated modelling approach comprised of multiple stages, from the data selection and filtering to the origin-to-destination estimation modelling across multiple modes, down to a dynamic assignment and microscopic simulation modelling aimed at evaluating the impact of disruptions across multiple modes. Finally, we propose a mode shift impact modelling to evaluate the best mitigation strategies and employ different disruption impact indicators such as delay time, flow, density and travel time for identifying the impacts.

Another important contribution represents the investigation of the mode shift behaviour according to dynamic traffic states; more specifically, we provide a method to examine the change in traffic states and the travel costs due to mode and route shifts under traffic disruptions. We model the decision-making en-route and the mode choice relies on an iterative traffic assignment; this means that, for those flexible travellers, the route choice is modifiable during their travelling, and the decision-making is more appealing than those travellers who are loyal to the initial routing plan. To summarise, the main theoretical and methodological contributions of this chapter are:

- an integrated OD estimation modelling framework for multi-modal transport networks,
- a suitable spatial-temporal disruption impact modelling via a multi-modal transport simulation approach,
- a dynamic traffic assignment model that simulates the mode shift behaviour via a dynamic demand adjustment,
- an analytic method regarding the mechanism of mode shift as well as its impact under traffic disruptions.

This chapter is organised as follows. In Section 5.2, the dynamic trip assignment model is discussed, and the details of the integrated OD estimation is highlighted in Section 5.2.2, followed by the methodology of traffic disruption modelling and the procedures for determining the spatial-temporal impact of traffic disruptions in Section 5.2.3. The model for mode shift and impacted trips identification are included in Section 5.2.4. The application of the proposed methods to a real network is presented in Section 5.3 and the results of the case study are demonstrated in Section 5.4. Finally, the research conclusion and the future directions are provided in Section 5.5.

5.2 Methodology

5.2.1 Modelling framework

Figure 5.1 showcases our proposed modelling framework for evaluating the impact of disruptions across multi-mode transport networks. The framework consists of three stages: at *Stage 0* we collect, filter and aggregate all the input data-sets (such as traffic flow counts, traffic control plans, incident logs, etc.); at *Stage 1* we propose a multi-modal demand estimation modelling with the purpose of obtaining an integrated multi-modal OD demand matrix (see details in Section 5.2.2); at *Stage 2* we further propose a dynamic trip assignment and demand refinement based on the impact of transport disruptions (as explained in Section 5.2.3), and finally, at *Stage 3* we construct various mode and route shift strategies and their impact on the traffic congestion, as further described in Section 5.2.4.

5.2.2 Integrated multi-modal OD estimation *(Stage 1)*

The multi-modal transport system is firstly modelled by implementing a four-step demand estimation model but adapted to multiple public transport modes, as shown in *Stage 1*- Figure 5.1, including trip generation and attraction *(Step 1)*, trip distribution *(Step 2)*, modal split *(Step 3)* and static trip assignment at the macroscopic level *(Step 4)*, followed by a calibrated travel demand path assignment plan *(Step 5)*. The study area consists of $Z_j,\ j = \{1 \dots J\}$ zones, and for each time period t, the travel demand matrix, which is the main output of this stage *(Output 5)*, is denoted as:

$$OD_k(t) = \left[T^k_{i,j}(t)\right]_{Z\times Z},\ i,j = \{1 \dots J\}, \tag{5.1}$$

where $T^k_{i,j}$ stands for the number of trips originating from zone i and arriving at zone j at time interval t by transport mode k. Due to space constraints, we provide the mathematical modelling of the Gravity-Model for multi-mode public transport in the supplementary material (see Online-supplement, 2022), while focusing in the following on the incident impact modelling.

5.2.3 Disruption modelling without mode shift (*Stage 2*)

By using the calibrated OD demand from *Stage 1* and the reported incident logs, we further apply a dynamic trip assignment which re-adjusts and generates a dynamic and time-dependent OD matrix that is used for: a) evaluating a baseline scenario where people travel as usual, without any disruptions (see *Output 6*), and b) evaluating the impact of reported accident logs but assuming that people do not make any changes to their trips, and instead wait for the incident to be cleared off (see *Output 7*).

Indicators for impact identification

The indicator we use when determining the impact of a traffic disruption D on link l during time period t is the ratio of the traffic state parameter υ, which is illustrated by the following formulas:

$$R^D_l(t) = \frac{\left(\upsilon^\alpha_l - \upsilon^{D,\beta}_l\right)(t)}{\upsilon^\alpha_l(t)} \tag{5.2}$$

$$\Delta v_l^D(t) = \left(v_l^\alpha - v_l^{D,\beta}\right)(t) \tag{5.3}$$

where v^α is the traffic state not affected by disruptions, $v_l^{D,\beta}$ is the traffic state affected by the disruption and $\Delta v_l^D(t)$ represents the scale of the impact; finally $R_l^D(t)$ stands for the ratio of a baseline versus a disrupted network. The total number of links in the study area is $l, l = \{1 \dots L\}$; α represents the baseline traffic situation; β represents the situation when the traffic disruption D is impacting the network at time period t.

To understand the change of traffic states, various indicators such as mean link delay time $t_l^{delay}(t)$, mean travel time $t_l^{travel}(t)$, flow $f_l^{flow}(t)$ or density $f_l^{density}(t)$ during time period t are used for ratio impact analysis provided later in the Results section.

The overall traffic disruption in the network that impacts the traffic states is represented by:

$$D(t) = g\left(\sum_{r=1}^{L} d_r^\gamma(t)\right) \tag{5.4}$$

where the impact of the disruption is the function of the sum of all appeared disruptions in the network during a time period t; d is the individual disruption event; r indicates the location of the disruption, $r = \{1 \dots L\}$; γ is the binary parameter that indicates whether the disruption is impacting or not the link l during t.

Temporal impact identification: The total impact duration (Δt_D^{impact}) of the network is the accumulation of the time period from the time when the travelling of a vehicle is first impacted by the disruption D ($t_D^{impact-initial}$) until the time when the first vehicle can travel as usual, without being impact, denoted as $t_D^{impact-end}$. Therefore, the total impact duration can be described by:

$$\Delta t_D^{impact} = t_D^{impact-end} - t_D^{impact-initial} \tag{5.5}$$

with additional constraints with regards to the initial time of the disruption ($t_D^{D-initial}$), the end of the disruption (t_D^{D-end}) and the parameters of traffic states ($v_l^{D,\beta}$ and v_l^α):

$$t_D^{impact-initial} \geq t_D^{D-initial} \tag{5.6}$$

$$\epsilon \sum_{l=1}^{L} v_l^{D,\beta}\left(t^{impact-initial}\right) \geq \sum_{r=1}^{L} v_l^{D,\alpha}\left(t^{impact-initial}\right) \tag{5.7}$$

$$t_D^{impact-end} \geq t_D^{D-end} \tag{5.8}$$

$$\epsilon \sum_{r=1}^{L} v_l^{D,\beta}\left(t^{impact-end}\right) \leq \epsilon \sum_{r=1}^{L} v_l^{D,\alpha}\left(t^{impact-end}\right) \tag{5.9}$$

where ϵ is the factor of impacted traffic state that indicates the pre-defined level of impact. For instance, if ϵ is 90% and the indicator is delay time, the link is assumed to be impacted by the disruption when 90% of the link delay time is greater than the delay time for the baseline scenario. As for the representations of traffic states, such as travel speed, that are reduced by the disruption, the associated constraints are switch from Equation 5.7 to:

$$\sum_{r=1}^{L} v_l^{D,\beta}\left(t^{impact-initial}\right) \leq \epsilon \sum_{r=1}^{L} v_l^{D,\alpha}\left(t^{impact-initial}\right) \tag{5.10}$$

from Equation 5.9 to:

$$\epsilon \sum_{r=1}^{L} v_l^{D,\beta}\left(t^{impact-initial}\right) \geq \sum_{r=1}^{L} v_l^{D,\alpha}\left(t^{impact-initial}\right) \tag{5.11}$$

Spatial impact identification via links: The time impact of the disruption is captured from a network-wide perspective, but the spatial impact can be analysed via a link-based analysis:

$$S^D(t) = [l^{\gamma}(t)] \tag{5.12}$$

where γ is the binary parameter that indicates whether the disruption is impacting the link l during t; similarly to the time impact, we define the following constraints when the impacted links can be identified if:

$$\epsilon v_l^{D,\beta}\left(t^{impact-initial}\right) \geq v_l^{D,\alpha}\left(t^{impact-initial}\right) \tag{5.13}$$

$$\epsilon v_l^{D,\beta}\left(t^{impact-end}\right) \leq \epsilon v_l^{D,\alpha}\left(t^{impact-end}\right) \tag{5.14}$$

The impact on the incident duration can be defined by using the affected links as follows:

$$\Delta t_{l,D}^{impact} = t_{l,D}^{impact-end} - t_{l,D}^{impact-initial} \tag{5.15}$$

Spatial impact identification via OD matrix: In terms of the zonal impact analysis, the affected number of trips from an OD pair can be determined by comparing the OD matrix with or without the disruption in the road network. The level of impact on trips is subjected to the ratio of the disparity of trips travelled between zones in a baseline versus incident scenario, which is derived from Equation 5.2 as follows:

$$R_{i,j}^{T}(t) = \frac{\left(T_{i,j}^{\alpha} - T_{i,j}^{\beta}\right)(t))}{\left(T_{i,j}^{\alpha} + \epsilon\right)(t)}, T_{i,j}^{\alpha}, T_{i,j}^{\beta} and \epsilon \geq \tag{5.16}$$

where $T_{i,j}^{\alpha}$ indicates the number of trips between zone i and j under a baseline scenario, $T_{i,j}^{D,\beta}$ represents the disruption-impacted number of trips and ϵ is a small constant that is added at the number of trips without impacting by any disruption in order to enable calculation, as in reality, some of the zone pairs are inaccessible for a specific transport mode. The different between the baseline and the disruption-impacted trips is denoted as $\Delta T_{i,j}(t)$. Therefore, if the ratio of disparity trip is between 0 and 1, then the number of trips inside the affected OD is reduced; while if the ratio is negative, the network experiences an increase in the affected number of trips which can be explained by alternative modes being deployed in the network by the affected number of people.

Disruption modelling

The impacted travel cost of every link during time t due to the reduction of link capacity should be described as a function of disruption duration (Equation 5.15) and disruption scale (Equation 5.12):

$$C_l(t) = h\,(t, l)\,, l(t) \in S^D(t), t \in \Delta t_D^{impact} \tag{5.17}$$

This means that, for each impacted link during the incident, the properties of the link can be highlighted by the capacity or by the limited travel speed change by time and location. Therefore, if we assume that the travel cost is majorly subjected to the travel speed, the link closure can be described as $C_l(t) = 0$, where the limited travel speed equals zero during the disruption period; in terms of the simulation of link capacity reduction, this can be achieved by introducing a weight to the designed limited travel speed V for each lane w of the link l, which is denoted by $\eta_{l,w}$. Therefore, the weighted limited travel speed can be illustrated by $\eta_{l,w}V_{l,w}$.

5.2.4 Disruption modelling under mode shift (*Stage 3*)

Mode shift modelling

The commonly used strategies to minimise travel costs are mode and route shifts. By mode shift, we refer to the shift of the travel demand between different means of transportation. Therefore, the objective of mode shift modelling is to first find the changeable demand for each transport mode, then adjust it accordingly in the travel demand matrix in order to evaluate its impact. Such process is what is illustrated as *Step 7* and *Step 8* with the highlighted *Outcome 8* in Figure 5.1.

Given a scenario of a road network disruption, the perceived travel cost increases, and the decision-making on mode and route choice is re-decided according to the utility function. The re-calculated set of the shortest path by a transport mode will update the travel demand. For the updated travel demand matrix, the number of impacted trips travelled by mode k_1 due to the disruption is transferred to demand travelled by mode k_2, the number of transferred trips equals $\delta T_{i,j}^{k_1,k_2}(t)$. Hence, the travel demand by a transport mode k_1 between each pair of zones is calibrated by the impacted demand:

$$T_{i,j}^{k_1}(t) = T_{i,j}^{k_1}(t) - \delta T_{i,j}^{k_1,k_2}(t) \tag{5.18}$$

With regards to a mode k_2, the number of zonal trips is increased by $\delta T_{i,j}(t)$, and is denoted as:

$$T_{i,j}^{k_2}(t) = T_{i,j}^{k_2}(t) + \delta T_{i,j}^{k_1,k_2}(t) \tag{5.19}$$

The origin travel demand matrices for transport modes k_1 and k_2 are denoted as $[T_{i,j}^{k_1}(t)]$ and $[T_{i,j}^{k_2}(t)]$, separately, while the mode-shift adjusted matrices can be illustrated by $[T_{i,j}^{k_1,\delta}(t)]$ and $[T_{i,j}^{k_2,\delta}(t)]$, for each. We make the observation that a similar approach can be undertaken regardless of the number of modes in the network.

Loyal and flexible travelling to route choice modelling

When disruptions occur, the change of travel cost and link properties will influence the shortest path searching at the process of trip assignment (see *Step 6* in Figure 5.1). This logic captures the reality that travellers value travel costs before or during their travelling, which is reflected by the frequency of decision-making on mode and route. Therefore, the travellers who have loyal travel behaviour tend to travel by following their initial shortest path, which is searched before departing; they will not make changes even if disruptions occur. However, travellers who are more sensitive to travel costs tend to re-update their shortest path more frequently. The shortest path searching frequency is denoted by λ, and for those travellers that are loyal to their daily route, λ equals the simulation period. This is means that the shortest path is only calculated once during the simulation period. For flexible travellers, the λ is set as 10 minutes, which means that the shortest path is re-adjusted every 10 minutes according to the updated traffic states. This concept is applied in *Scenario 3* and further detailed in Section 5.4.3. This decision-making of mode and route can be simulated by a discrete choice model such as the multinomial logit model (ML) through a function of utilities of all path alternatives. The impact of a travel cost change on decision-making can also be well illustrated by applying a scale parameter to the utility in the multinomial logit model; therefore, the ML model is adopted in this research study.

5.3 Case study

5.3.1 Geography information

The case study model is implemented in the Aimsun software and represents the city of Tarragona in Spain, which contains 15 centroids, 71 nodes and 201 links. The nodes consist of 13 signalled nodes (signalised intersections) and 58 connection nodes (unsignalised intersections or turns). The links contain 201 road sections, including primary streets, freeways, ramps, roads, roundabout streets, secondary streets, urban roads, tertiary streets and toll roads with different link capacities and pre-defined limited travel speeds. There are 15 bus lines serving this area with 38 public transport stops. The land use data along with the model is collected from the census in 2012, and the initial travel cost data is generated from the trip matrices united in the vehicle for car and public transport users.

5.3.2 Hypothesised disruption details

To showcase our approach, we study the impact of the traffic disruption that took place at road section 300, as highlighted by the red shape in Figure 5.2.

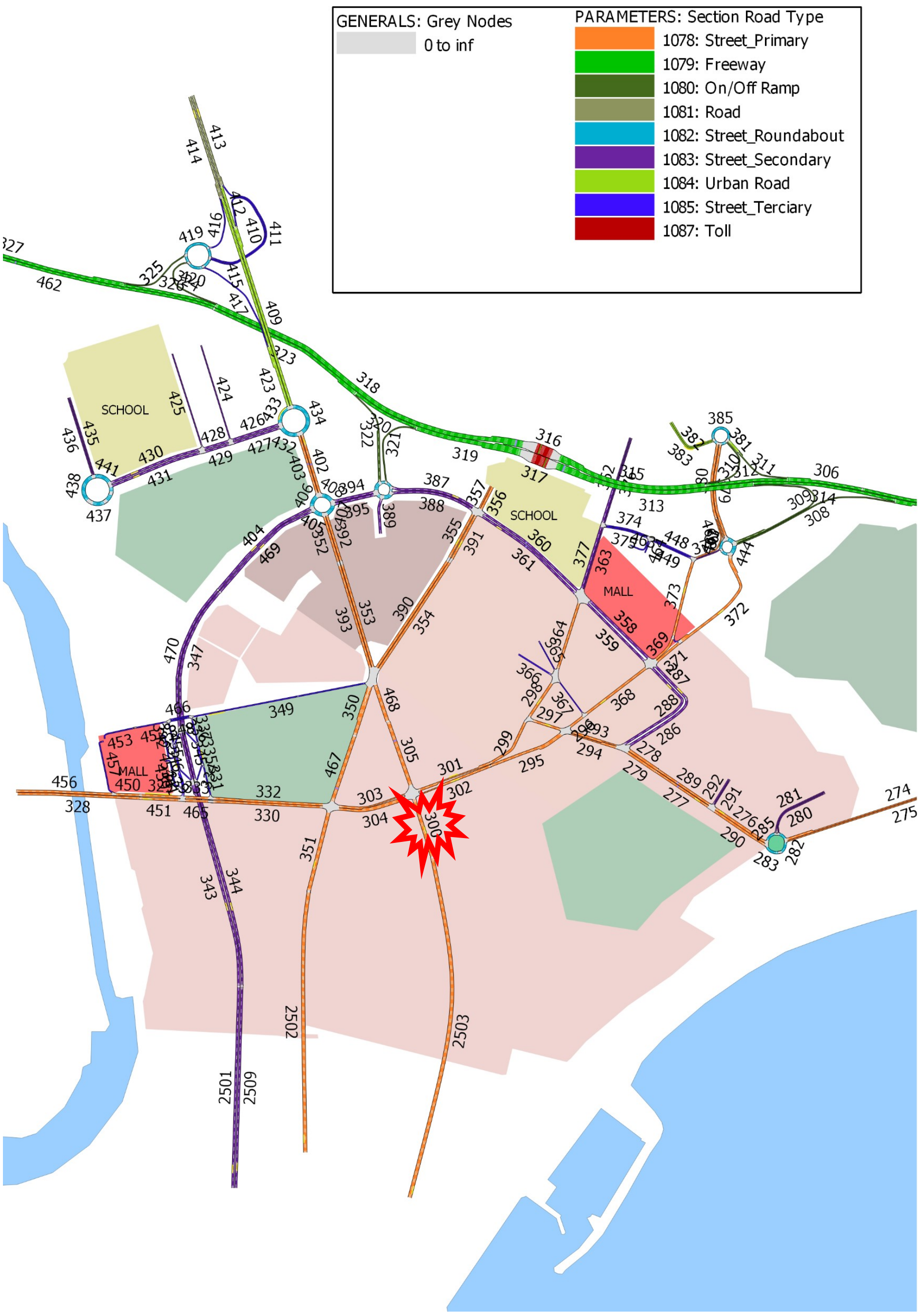


FIGURE 5.2: Map of the Tarragona area showing road networks and the disruption location.

The definition and characteristics of the modelled disruptions are further categorised and described by the following scenarios, where *Scenario 1* and *Scenario 2* intend to explore the consequences of the disruption in time and space, while *Scenario 3* attempts to understand the impact of mode shift on the network performance as well as the network mobility. The setting details of each scenario and experiment are presented in **??** In **Scenario 1**, we conduct three experiments categorised by the disruption duration which can last 10, 30 or 60 minutes, according to the Equation 5.17 and the explanation provided in Section 5.2.3 regarding the link closure. Any vehicle impacted by the disruption drops the travel speed to 0 km/h at the section. These experiments are named as *whole lane suspended for 10,* *30 and 60 minutes* in the next sections.

Scenario 2 contains a further three experiments defined by the disruption scale, when only one lane is impacted by the incident for 30 minutes while the rest of the lanes on the affected section remain functional, but the travel speed is reduced by 5, 25 and 30 km/h from 8:00 to 8:30 AM for each experiment. Such scenarios describe the link capacity reduction as expressed in Section 5.2.3. These experiments are named as *single lane speed drop by 5, 25 and 30 km/h* in the next sections.

In **Scenario 3**, we block the entire section from 8:00 to 8:30, and the travel speed drops to zero km/h during the disruption, where the mode choice ahead or after represents the situation when travellers are notified of the disruption before departing or after being blocked by a disruption; the loyal or flexible route choice is represented by the frequency of the shortest path searching during the simulating period. There are four experiments included in this scenario:

- *Experiment 1 (S3E1):* considers the situation when travellers are notified of the disruption before departing and make the decision of their mode choice ahead of the travelling, then choose their regular transport mode for the entire simulation. The new mode choice influenced by the disruption is identified by using the "shifted ahead" travel demand at the beginning of the simulation (this means that the travel utility calculation for the shortest path searching only happens once at the beginning of the simulation). These experiments are named as *mode shift ahead, loyal to route choice* in the following subsections;

- *Experiment 2 (S3E2):* expresses the circumstances that, before the disruption, travellers follow the baseline demand and travel loyally according to their normal route choice; however, after the disruption occurs, they choose a new mode choice for finishing their trips. The travel utility for the shortest path searching is calculated twice: at the beginning of the simulation and after the shifted demand is applied at 8:30 AM (after the disruption occurrence) in order to simulate the updated route choice for passengers. These experiments are further referred to as *mode shift after, loyal to route choice*;

- *Experiment 3 (S3E3):* illustrates the same setups as for *Experiment 1* but instead, we are searching for the best path every 10 minutes to simulate a flexible travel behaviour for travellers. This experiment is referred to as*mode shift ahead, flexible to route choice* in the following subsections;

- *Experiment 4 (S3E4)*: demonstrates the setups that include both a mode and a flexible route choice (meaning all travellers can switch between any type of mode and also between their routing to finish their destination). This experiment closely follows real-life behaviour and is referred to as *mode shift after, flexible to route choice* in the following sections.

For all modelled experiments in all three scenarios, the blockage clear time is estimated based on each vehicle following the rule that the vehicle departs immediately at the end of the blockage, which is the end time of disruption ($t^{impact-end}$). The earlier-arrived vehicles depart ahead of the later-blocked vehicles.

5.4 Results

5.4.1 Scenario 1 results: impact of various disruption duration

The results of modelling the impact of various disruptions according to *Scenario 1* are shown in Figure 5.4 and Figure 5.5. These two figures demonstrate the time-dependent change of traffic states, namely delay time, flow and density, using a link closure as the disruption modelling method (defined in Section 5.2.3). The traffic state ratios shown in Figure 5.5 are calculated by using Equation 5.2. The results are compared to the *Baseline* states, which allows us to investigate the impacts of various disruptions from the baseline.

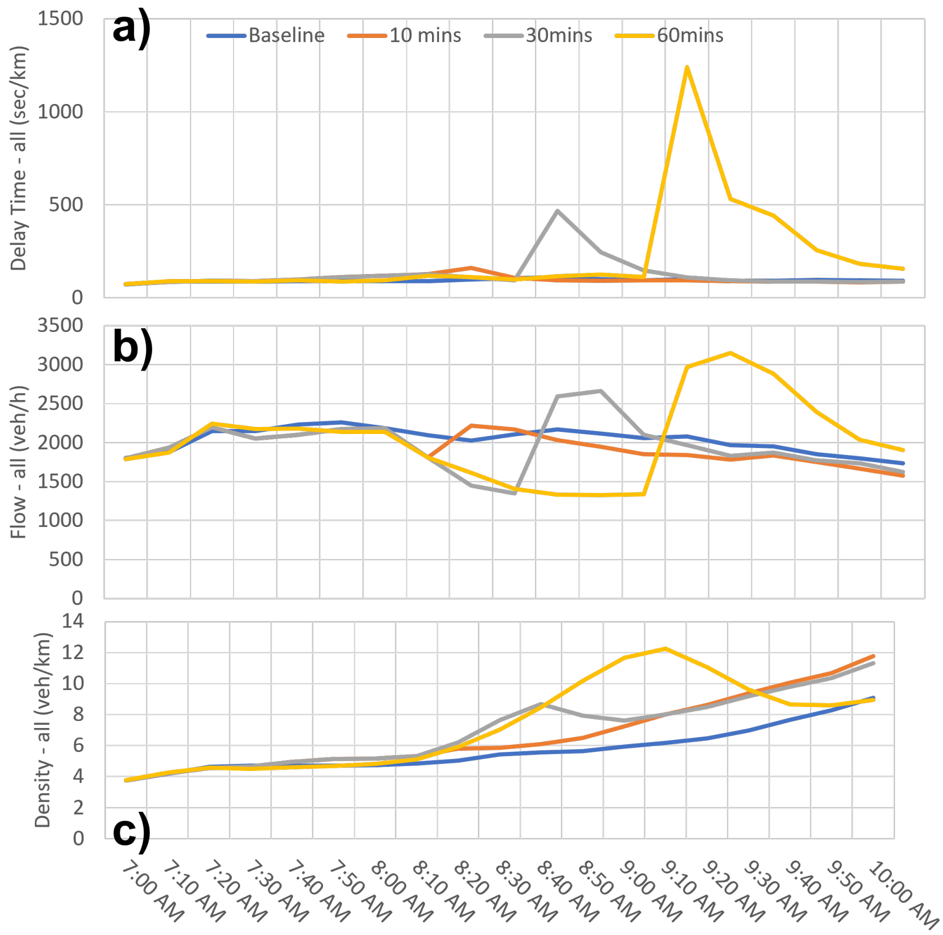


FIGURE 5.4: Impact of various disruption duration on a) delay time, b) flow and c) density

Figure 5.4 and Figure 5.5 combines the outputs for the morning peak hours, where the disruptions occur at 8:00 AM and last for 10, 30 or 60 mins. The mean delay time shown in Figure 5.4a) shows a significant impact post-accident, especially after 08:30-9 AM. When comparing the delay time of the baseline and their impact ratios (Figure 5.4a and Figure 5.5a) we observe that: a) for a 10-min disruption ending at 08:20 AM, there is a delay time ratio of -63%, b) for a 30-min disruption occurring at 08:40, the delay ration quickly reaches -312% (4.59 times higher than that of a small incident); and this is more severe c) for a 60-min disruption reaching its delay peak at 09:10 AM, when the delay ration is -1128 (more than ten times from the baseline).

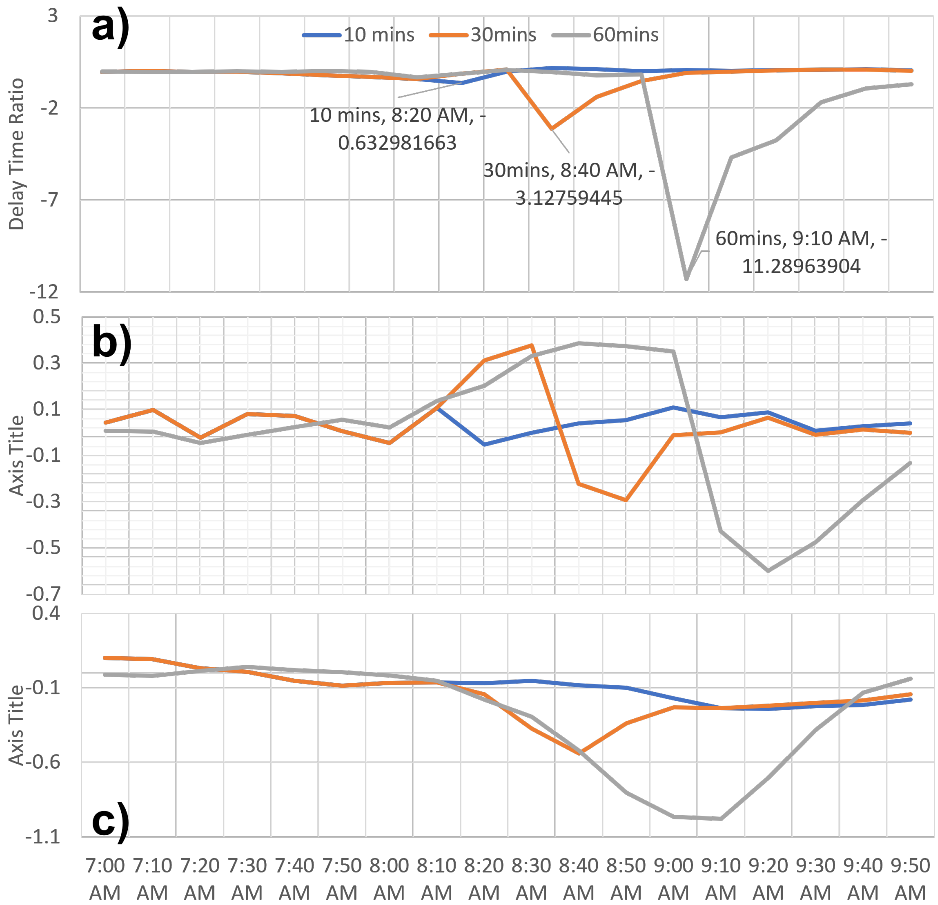


FIGURE 5.5: Impact of various disruption duration on a) delay time ratio, b) flow ratio and c) density ratio

5.4.2 Scenario 2 results: impact of various disruption scales

The temporal impact of a disruption is also highly related to the disruption speed scale, as shown in Figure 5.6, where the curves of the delay time, flow and density ratios are presented. Following the *Scenario 2* details from Section 5.3.2 and the disruption modelling method externalised by a capacity reduction in Section 5.2.3, the result indicates that the severity of the disruption against delay time increases with the single lane speed drop. From Figure 5.6a), we can see that the delay time does not change much in the three experiments, though there is still a peak in delay time ratio which appears at the same time (around 8:10 AM). This means that the scales of disruption could hard influence the occurrence of the peak delay. Figure 5.6b) and Figure 5.6c) show that the hypothesised single-lane speed drop also has a slight negative impact on the traffic flow and density. As shown by curves, a speed drop by 30 km/h (orange curve) results in more change in flow but it rarely impacts the change in density.

5.4.3 Scenario 3 results: impact of mode and route shift

The impact of mode shift is analysed according to the *Scenario 3* defined in Section 5.3.2, where the main strategy is a travel demand adjustment: affected passengers switch from the impacted public transport to cars following either a loyal and/or flexible route choice behaviour. At this stage, for

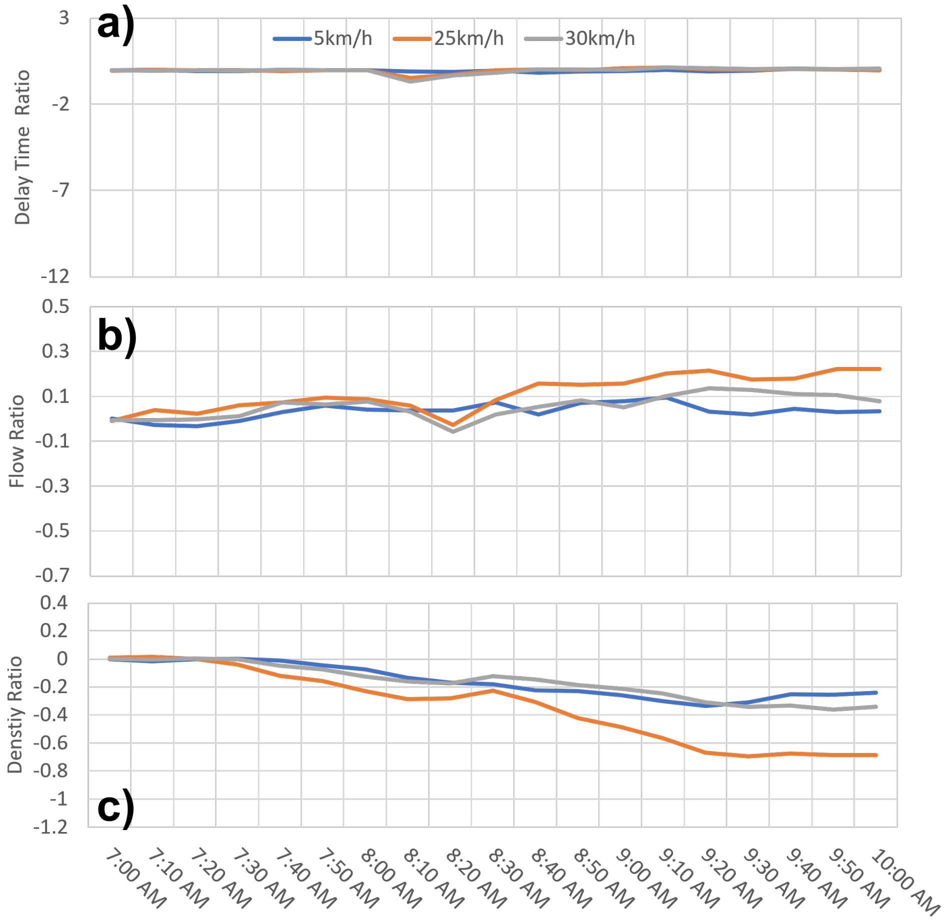


FIGURE 5.6: Impact of various disruption scales on a) delay time ratio, b) flow ratio and c) density ratio

those experiments applying mode shift ahead of departing (*S3E1, S3E3*), the adjusted public transport demand is increased while the car demand is decreased from the beginning of the simulation. While for the experiments that apply the mode shift after disruption (*S3E2, S3E4*), the demand adjustments are conducted after the disruption ends. The loyal and flexible route choice behaviours have been depicted in Section 5.2.4.

The shifting on travel demand (*Step 7*) is illustrated in Figure 5.1 and after another round of dynamic trip assignments based on the shifted dynamic travel demands via the four scenarios (*Step 8-11*), we analyse the *Outputs 8-11* which reflect the impact of mode shift and route shift on the road network performances, as shown in Figure 5.7 and Figure 5.8.

Figure 5.7 provides a comparison of the estimated travel demand across morning peak hours (7:00-10:00 AM), for the following experiments: *Scenario 1 Experiment 2 (S1E2: whole lane suspended for 30 minutes)*, *Scenario 3 Experiment 1 (S3E1: mode shift ahead, loyal to route choice)*, *Scenario 3 Experiment 2 (S3E2: mode shift after, loyal to route choice)*, *Scenario 3 Experiment 3 (S3E3: mode shift ahead, flexible to route choice)* and *Scenario 3 Experiment 4 (S3E4: mode shift after, flexible to route choice)*.

As shown in Figure 5.7, the mode share maintains the same for experiments *Baseline* and *S1E2*, while after the demand adjustment ahead of travelling (see results of *S3E1* and *S3E3*), the shares of car increase more dramatically than the demand adjustment after disruption (see results of *S3E2* and

S3E4) when the demands for public transport decrease contrarily. The impacts of mode and route shift on traffic states are further discussed in following Section 5.4.3 and Section 5.4.3.

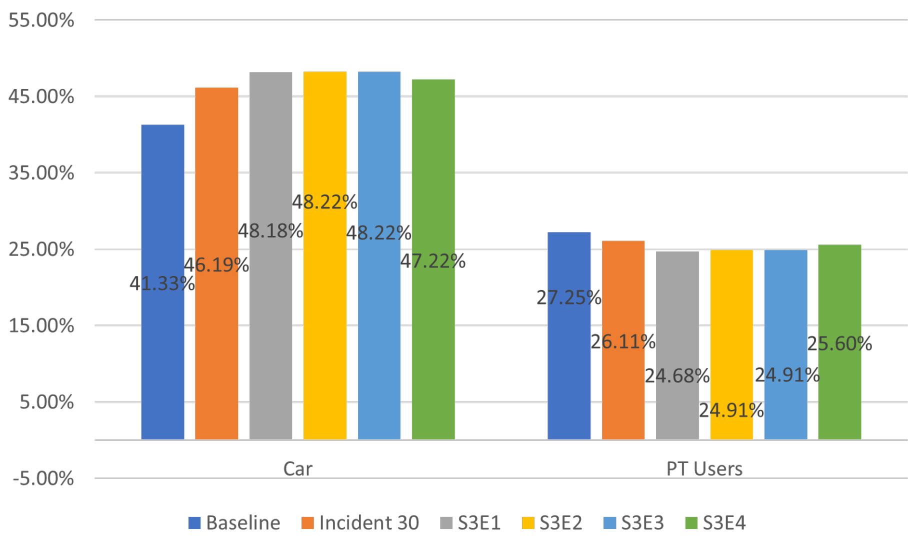


FIGURE 5.7: Comparison of travel distribution by transport modes

Impact of mode shift

The impact of mode shift is compared by analysing the delay time ratio $R((t_l^{delay}(t))$, the flow ratio $R(f_l^{flow}(t))$ and the density ratio $R(f_l^{density}(t))$, based on experiments *S3E1* and *S3E2*. As shown in Figure 5.8 a), both modes shift ahead and after the disruption benefit the travellers by reducing delay time, and the mode shift ahead (*S3E4*) is slightly superior to mode shift after (*S3E2*) as can be observed by analysing the timing around 8:50 to 9:30 in this case. By looking at the curves of flow ratio (Figure 5.8 b)) and density ratio (Figure 5.8 c)), both mode shift strategies increase travel flow after disruption and ease the traffic congestion. The curves of flow ratio indicate that before and during the disruption, the mode shift ahead increases the flow, but due to the extra demand, the performance of the post-disruption congestion relief is limited. An interesting phenomenon is that, right after the disruption, the road section is free of the vehicle; this enables the initially released vehicles to run in a free-flow situation, which results in a peak of flow that is even higher than the *Baseline* situation. The density curves in Figure 5.8c) indicate that shifting modes after the disruptions is more efficient on easing the congestion because the bottom density ratio is reduced only to -24% for *S3E2* (see grey curve) while for *S3E1* it can only touch down to -33%, and the bottom density ratio is -34% for *S3E3* (yellow curve) while for *S3E4* it is -39%, comparing with the ratio of *Incident 30* (dark blue curve) is -56%.

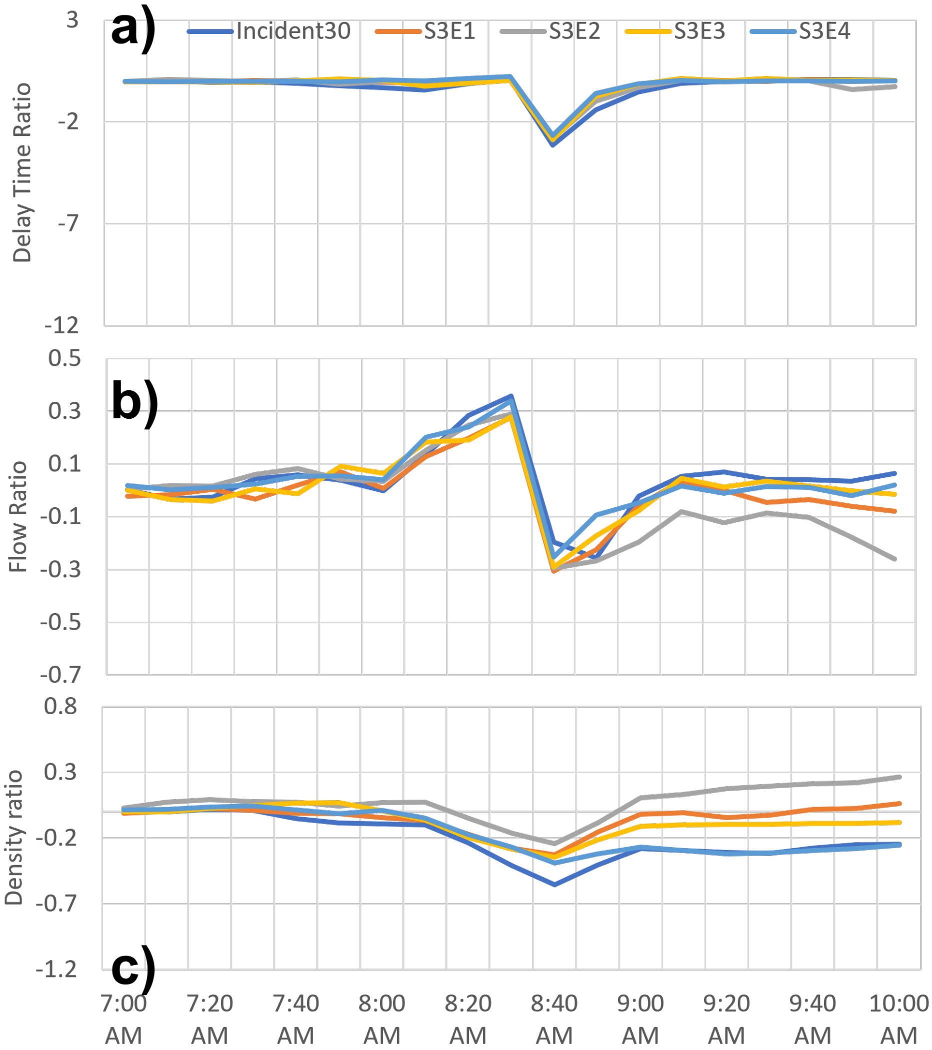


FIGURE 5.8: Comparison of the impact of mode and route choice on a) delay time ratio, b) flow ratio and c) density ratio

Impact of route shift

The route shift is modelled using a stochastic route choice, where trips can be assigned to the network after the shortest travel path for each trip, as introduced in Section 5.2.4. To understand the impact of route shift on network performance, results based on *S3E1* verses *S3E3*, and *S3E2* verses *S3E4* are selected accordingly. The significant benefit in delay time reduction can be observed when applying the *S3E4* with mode shift applied after disruptions and flexible route choice behaviour, where the delay time ratio is reduced to -267% and the ratio of flow is reduced to -25% (see Figure 5.8 a) and b)). According to Figure 5.8 c), the higher density ratio appears on the curve of *S3E2*(-24%), which indicates that the strategy with flexible route choice with mode shift ahead performs much more gratifying.

5.5 Conclusion

This chapter proposed an integrated multi-modal hybrid modelling framework that embeds the four-step model estimation with a dynamic demand estimation and mode shift approach. This framework demonstrates the potential ability to model real-time disruptions, their impact and an evaluation of mode shift and route shift in a multi-modal environment. We consider the essentials of dynamic microscopic transport simulation and propose a methodology to identify and quantify the temporal-spatial impact of transport disruption. For the spatial impact analysis, we proposed a method to detect the impacted trips between OD pairs. This permits us to simulate the mode choice behaviour without starting from the static demand estimation, which, therefore, does not require more land-use data. The results and findings generated from this research study evidence that public transport does make travel more accessible, especially when disruption is suspending the traffic. The mode and route shift benefit the transport system by increasing the flow and effectively reducing delay time and density. Future investigations could cover the topic of examining more types of disruptions by duration or location by links. The theoretical link-based disruption impact estimation method is proposed in this chapter, but the chapter falls under statistical analysis following this idea. More investigation towards impact propagation into the network through links could be a good sub-direction. The impact of dual or multiple disruptions on network performance is also underestimated in this chapter. We limit the work only by considering the single disruption, single land closure and single road section closure. The method to quantify the impact of multiple disruptions in the network is also required. Investigating data-driven methods for impact measurement and model calibration is another highly recommended direction.

Chapter 6

Conclusion

6.1 Conclusion of findings and contributions

This research study first introduces a novel framework for dynamic large-scale OD estimation in PT systems, as detailed in Chapter 3. The framework focuses on a microscopic stop-based OD matrix and employs a 15-minute time interval to simplify computations while mimicking dynamic conditions. By calculating trips between OD pairs every 15 minutes, the model assesses the effects of various travel cost matrices, including single, multiple, and entropy-weighted features, using metrics like MAE, RMSE and MAPE. The framework is based on the Gravity Model with inputs such as smart card data for generating attraction vectors and PT GTFS data for cost matrix features.

A novel deterrence function calibration method utilizing Shannon's entropy is proposed, pre-weighting cost features before iterative parameter calibration, thus reducing computational load. The performance of this method is compared with traditional Hyman's and Traverse Searching methods, with Traverse Searching showing promise in identifying optimal parameters for bus networks. The fusion of travel cost features from Hyman's and Entropy-weighted methods further enhances accuracy, particularly with an optimal combination of costs like closeness and straightness. While fusion decreases accuracy initially, effective combinations of practical cost features lead to notable accuracy improvements.

Moreover, analysis of mean errors over time reveals fluctuations correlating with network trip numbers, indicating that network overcrowding influences OD estimation accuracy. This highlights the model's ability to capture real-time dynamics and its sensitivity to network congestion levels.

With the accurate estimated OD matrix, we are able to process to the next step, which is incident impact identification and modelling. In the following Chapter 4, we present a method to dynamically model PT patronage patterns and analyze the impacts of road incidents on PT users. Our proposed method utilizes the Fourier transform to decompose intricate patterns into distinct waves, allowing us to capture dominant components of patronage patterns and use them as reference profiles for traffic analysis. Specifically, we focus on incident impact identification, a challenging task due to peak hours' influence on accurately assessing current incident impacts. To address this challenge, we conduct an enhanced sensitivity test to improve the modelling ability for typical days, thereby effectively capturing current impacts. Multiple sample incidents are tested using this method, demonstrating robust results.

While our modelling method allows for the identification of optimal typical profiles, its effectiveness relies on substantial count data and specific disruptions, such as prolonged durations and locations distant from PT-only lanes, to generate precise results. These limitations present opportunities for future research enhancements. We propose focusing on nearby stop data analysis, which could significantly enhance the model's ability to identify and assess impacts. Additionally, future directions may involve decomposing patterns and effectively identifying noise generated by disruptions, such as recurring congestion or incidents, to quantify their impacts using wave functions.

Moreover, our model holds the potential for expansion to incorporate spatial analysis, introducing an additional spatial dimension to capture the evolution of impacts based on location. This extension would offer valuable insights into the spatial dynamics of disruptions and their effects on PT patronage patterns. Overall, our method provides a framework for dynamic PT patronage modelling and incident impact identification, with the potential for further refinement and expansion to enhance its effectiveness in real-world applications.

Chapter 5 introduces an innovative integrated multi-modal hybrid modelling framework, merging the four-step model estimation with dynamic demand estimation and mode shift strategies. This framework showcases promising potential in modelling real-time disruptions, and their consequences, and evaluating mode and route shifts within a multi-modal context. Acknowledging the importance of dynamic microscopic transport simulation, we propose a methodology to precisely identify and quantify the temporal-spatial impact of transport disruptions. Our spatial impact analysis introduces a method to identify affected trips between OD pairs, allowing simulation of mode choice behaviour without relying on static demand estimation or extensive land-use data. This part of the research provides compelling evidence that public transport significantly enhances travel accessibility, particularly during traffic disruptions. Mode and route shifts prove beneficial for the transport system, facilitating increased flow and effectively reducing delay time and density. By integrating dynamic demand estimation and mode shift approaches, our framework offers a comprehensive understanding of the complex dynamics within multi-modal transportation systems.

6.2 Conclusion of future directions

In our research, we first present a framework for dynamic stop-by-stop OD matrix estimation tailored for large-scale PT networks. However, our model currently overlooks the impact of general traffic factors, such as delay time, which could be addressed by comparing scheduled travel times with historical data from smart-card records.

While our proposed PT OD matrix estimation model is initially applied to bus and train networks due to data availability, it holds potential for integration into other transport modes like cars, light-rails, or on-demand solutions to create a comprehensive multi-modal OD matrix estimation model. Future research could explore transfer patterns between bus and train networks, aiming to refine the modelling process and better capture transfer trips.

Furthermore, there's scope for further exploration of deterrence function calibration parameters. Although we have focused on the Hyman and Traverse Searching methods in this chapter, numerous other estimation models and deterrence function forms warrant investigation. Leveraging smart-card data opens avenues for testing various methods like regression models, maximum likelihood estimation, Bayesian estimation, and machine learning algorithms to generate OD matrices based on historical data.

Examining traverse search results reveals a tipping point where MAE stabilizes despite parameter variations, suggesting potential avenues for understanding this phenomenon in future research. Additionally, investigating the tipping point concerning travel cost features not explored in this study could shed light on the relationship between travel cost and OD matrix estimation accuracy, providing valuable insights for improving modelling approaches.

The proposed impact modelling method (see Chapter 4) enables the identification of an optimal typical profile, yet its effectiveness relies on a significant amount of count data and specific disruptions. These limitations suggest opportunities for future research enhancements. Focusing on nearby stop data could notably improve the model's impact assessment capability. Future directions may involve decomposing patterns to identify disruptive noise, facilitating quantification of disruption impacts using wave functions. Additionally, expanding the model to include spatial analysis could capture the evolving impacts based on location, offering valuable insights into spatial dynamics and their effects on public transport patronage patterns.

In terms of the journey recovery for PT users under disruption, future investigations should explore various types of disruptions based on duration or location, focusing on link-level impact estimation within the transportation network. While this chapter introduces a theoretical link-based disruption impact estimation method, it primarily relies on statistical analysis. Further research could delve deeper into understanding how disruptions propagate through network links. Additionally, the chapter primarily addresses single disruptions, such as land closures or road section closures, overlooking the impact of dual or multiple disruptions on network performance. It is crucial to develop methods to quantify the cumulative impact of multiple disruptions within the network. Moreover, exploring data-driven approaches for impact measurement and model calibration is highly recommended for future research endeavours in this field.

Bibliography

Yi, Zhiyan et al. (2021). “Inferencing hourly traffic volume using data-driven machine learning and graph theory”. In: *Computers, Environment and Urban Systems* 85, p. 101548. ISSN: 0198-9715. DOI: 10.1016/J.COMPENVURBSYS.2020.101548.

PastorSatorras, R. and M. Rubi A. Diaz, Guilera (2003). *Statistical Mechanics of Complex Networks*. Vol. 625. Lecture Notes in Physics. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-540-40372-2. DOI: 10.1007/b12331. URL: http://link.springer.com/10.1007/b12331.

Clauset, Aaron, Cosma Rohilla Shalizi, and M. E.J. Newman (2009). “Power-Law Distributions in Empirical Data”. In: *http://dx.doi.org/10.1137/070710111* 51.4, pp. 661–703. ISSN: 00361445. DOI: 10.1137/070710111. arXiv: 0706.1062. URL: https://epubs.siam.org/doi/abs/10.1137/070710111.

Ortuzar S., Juan de Dios. and Luis G. Willumsen (2011). *Modelling transport*. WILEY, p. 607. ISBN: 9780470760390. DOI: 10.1002/9781119993308. arXiv: arXiv:1011.1669v3. URL: https://onlinelibrary-wiley-com.ezproxy.library.sydney.edu.au/doi/book/10.1002/9781119993308.

Shen, Yi, Gang Ren, and Bin Ran (2019). “Cascading failure analysis and robustness optimization of metro networks based on coupled map lattices: a case study of Nanjing, China”. In: *Transportation* 48.2, pp. 537–553. ISSN: 1572-9435. DOI: 10.1007/S11116-019-10066-Y. URL: https://link.springer.com/article/10.1007/s11116-019-10066-y.

Shang, Wen Long et al. (2020). “Statistical Characteristics and Community Analysis of Urban Road Networks”. In: *Complexity* 2020. ISSN: 10990526. DOI: 10.1155/2020/6025821. arXiv: 1907.12367.

Saberi, Meead et al. (2020). “A simple contagion process describes spreading of traffic jams in urban networks”. In: *Nature Communications* 11.1, pp. 1–9. ISSN: 2041-1723. DOI: 10.1038/s41467-020-15353-2. arXiv: 1906.00585. URL: https://www.nature.com/articles/s41467-020-15353-2.

Zhao, Dong et al. (2022). “Traffic disruption modelling with mode shift in multi-modal networks”. In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*. Vol. 2022-Octob. Institute of Electrical and Electronics Engineers Inc., pp. 2428–2435. ISBN: 9781665468800. DOI: 10.1109/ITSC55140.2022.9921763. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85141834209&doi=10.1109%2FITSC55140.2022.9921763&partnerID=40&md5=3fb0baae8c4f0218eab281da45e1cd76.

Department of Infrastructure and Regional Development (2016). *Transport and Australia’s Development to 2040 and Beyond*. Commonwealth of Australia, Commonwealth, pp. 1–52. ISBN: 9781925401813. URL: https://www.infrastructure.gov.au/infrastructure/publications/files/Trends_to_2040.pdf.

OECD (2018). *Key Transport Statistics 2018 (2017 Data) | ITF*. URL: https://www.itf-oecd.org/key-transport-statistics-2018-2017-data (visited on 11/01/2020).

Mitchell, Robert B. and Chester Rapkin (1954). "Urban Traffic". In: *Urban Traffic*, p. 226. DOI: 10.7312/MITC94522/HTML.

Page, Matthew J. et al. (2021). "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews". In: *The BMJ* 372. DOI: 10.1136/BMJ.N71.

Renyi, Erdos (1959). "On random graph". In: *Publicationes Mathematicate* 6.6, pp. 209–297. URL: https://ci.nii.ac.jp/naid/10026207309/.

Watts, Duncan J. and Steven H. Strogatz (1998). "Collective dynamics of 'small-world' networks". In: *Nature* 393.6684, pp. 440–442. ISSN: 00280836. DOI: 10.1038/30918. URL: http://us.imdb.comhttps://www-nature-com.ezproxy.library.sydney.edu.au/articles/30918.

Barabási, Albert László and Réka Albert (1999). "Emergence of scaling in random networks". In: *Science* 286.5439, pp. 509–512. ISSN: 00368075. DOI: 10.1126/science.286.5439.509. arXiv: 9910332 [cond-mat]. URL: www.sciencemag.orghttp://science.sciencemag.org/.

Lin, Jingyi and Yifang Ban (2013). "Complex Network Topology of Transportation Systems". In: *Transport Reviews* 33.6, pp. 658–685. ISSN: 01441647. DOI: 10.1080/01441647.2013.848955. URL: https://www.tandfonline.com/doi/abs/10.1080/01441647.2013.848955http://www.tandfonline.com/doi/abs/10.1080/01441647.2013.848955.

Shanmukhappa, Tanuja et al. (2019). "Recent Development in Public Transport Network Analysis from the Complex Network Perspective". In: *IEEE Circuits and Systems Magazine* 19.4, pp. 39–65. DOI: 10.1109/MCAS.2019.2945211.

Regt, Robin de et al. (2018). "Public transportation in Great Britain viewed as a complex network". In: *https://doi.org/10.1080/23249935.2018.1529837* 15.2, pp. 722–748. ISSN: 23249943. DOI: 10.1080/23249935.2018.1529837. URL: https://www.tandfonline.com/doi/abs/10.1080/23249935.2018.1529837.

Dimitrov, Stavri Dimitri and Avishai Ceder (2016). "A method of examining the structure and topological properties of public-transport networks". In: *Physica A: Statistical Mechanics and its Applications* 451, pp. 373–387. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2016.01.060.

Shanmukhappa, Tanuja, Ivan Wang-hei Ho, and Chi Kong (2018). "Spatial analysis of bus transport networks using network theory". In: *Physica A* 502, pp. 295–314. ISSN: 0378-4371. DOI: 10.1016/j.physa.2018.02.111. URL: https://doi.org/10.1016/j.physa.2018.02.111.

Luo, Ding et al. (2019). "Integrating network science and public transport accessibility analysis for comparative assessment". In: *Journal of Transport Geography* 80, p. 102505. ISSN: 0966-6923. DOI: 10.1016/J.JTRANGEO.2019.102505.

Du, Zhouyang et al. (2020). "Identifying critical nodes in metro network considering topological potential: A case study in Shenzhen city—China". In: *Physica A: Statistical Mechanics and its Applications* 539, p. 122926. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2019.122926.

Huynh, Nam and Johan Barthelemy (2021). "A comparative study of topological analysis and temporal network analysis of a public transport system". In: *International Journal of Transportation Science and Technology*. ISSN: 2046-0430. DOI: 10.1016/J.IJTST.2021.05.003.

Zhang, Yifan and S. Thomas Ng (2021). “A hypothesis-driven framework for resilience analysis of public transport network under compound failure scenarios”. In: *International Journal of Critical Infrastructure Protection* 35, p. 100455. ISSN: 1874-5482. DOI: 10.1016/J.IJCIP.2021.100455.

Kopsidas, Athanasios and Konstantinos Kepaptsoglou (2022). “Identification of critical stations in a Metro System: A substitute complex network analysis”. In: *Physica A: Statistical Mechanics and its Applications* 596, p. 127123. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2022.127123.

Yang, Zhaosheng et al. (2014). “Research on short-term traffic flow prediction method based on similarity search of time series”. In: *Mathematical Problems in Engineering* 2014. ISSN: 15635147. DOI: 10.1155/2014/184632.

De Bona, A. A. et al. (2016). “Analysis of public bus transportation of a Brazilian city based on the theory of complex networks using the P-space”. In: *Mathematical Problems in Engineering* 2016. DOI: 10.1155/2016/3898762.

Zhang, Pei Pei et al. (2006). “Model and empirical study on some collaboration networks”. In: *Physica A: Statistical Mechanics and its Applications* 360.2, pp. 599–616. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2005.05.044.

Sun, Lijun, Yang Lu, and Der-Horng Lee (2015). “Understanding the Structure of Urban Bus Networks: The C-Space Representation Approach”. In: *CICTP 2015 - Efficient, Safe, and Green Multimodal Transportation - Proceedings of the 15th COTA International Conference of Transportation Professionals*, pp. 1557–1567. DOI: 10.1061/9780784479292.143. URL: https://ascelibrary.org/doi/abs/10.1061/9780784479292.143.

Gallotti, Riccardo and Marc Barthelemy (2015). “The multilayer temporal network of public transport in Great Britain”. In: *Scientific Data 2015 2:1* 2.1, pp. 1–8. ISSN: 2052-4463. DOI: 10.1038/sdata.2014.56. arXiv: 1501.02159. URL: https://www.nature.com/articles/sdata201456.

Feng, Jia et al. (2017). “Weighted complex network analysis of the Beijing subway system: Train and passenger flows”. In: *Physica A: Statistical Mechanics and its Applications* 474, pp. 213–223. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2017.01.085.

Yildirimoglu, Mehmet and Jiwon Kim (2018). “Identification of communities in urban mobility networks using multi-layer graphs of network traffic”. In: *Transportation Research Part C: Emerging Technologies* 89, pp. 254–267. ISSN: 0968-090X. DOI: 10.1016/J.TRC.2018.02.015.

Sui, Yi et al. (2019). “Public transport network model based on layer operations”. In: *Physica A: Statistical Mechanics and its Applications* 523, pp. 984–995. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2019.04.269.

Tang, Jinjun et al. (2021a). “Identifying critical metro stations in multiplex network based on D-S D–S evidence theory”. In: *Physica A: Statistical Mechanics and its Applications* 574, p. 126018. ISSN: 03784371. DOI: 10.1016/j.physa.2021.126018.

Yang, Haihua and Shi An (2021). “Robustness evaluation for multi-subnet composited complex network of urban public transport”. In: *Alexandria Engineering Journal* 60.2, pp. 2065–2074. ISSN: 1110-0168. DOI: 10.1016/J.AEJ.2020.12.016.

Liu, Zhizhen et al. (2022b). “Evaluating the dynamic resilience of the multi-mode public transit network for sustainable transport”. In: *Journal of Cleaner Production* 348, p. 131350. ISSN: 09596526. DOI: 10.1016/j.jclepro.2022.131350.

Pu, Han, Yinzhen Li, and Changxi Ma (2022). “Topology analysis of Lanzhou public transport network based on double-layer complex network theory”. In: *Physica A: Statistical Mechanics and its Applications* 592, p. 126694. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2021.126694.

Schoemaker, Th.J.H., K Koolstra, and P.H.L. Bovy (1999). *Traffic in the 21st century - a scenario analysis for the traffic market in 2030*. Tech. rep. Delft SCHÖNHARTING: Delft University Press, pp. 175–194.

Binsbergen, Arjan Van and Johan Visser (2001). “Innovation Steps Towards Efficient Goods Distribution Systems for Urban Areas”. PhD thesis. Delft University of Technology, p. 576. ISBN: 9040721793.

Nes, Rob van (2002). “Design of multimodal transport networks A hierarchical approach”. PhD thesis. Delft University of Technology, p. 441. ISBN: 9040723141. DOI: 10.1145/564437.564484.

Kivelä, Mikko et al. (2014). “Multilayer networks”. In: *Journal of Complex Networks* 2.3, pp. 203–271. ISSN: 20511329. DOI: 10.1093/comnet/cnu016. arXiv: 1309.7233. URL: https://academic.oup.com/comnet/article/2/3/203/2841130.

Boccaletti, S. et al. (2014). “The structure and dynamics of multilayer networks”. In: *Physics Reports* 544.1, pp. 1–122. DOI: 10.1016/J.PHYSREP.2014.07.001.

Aleta, Alberto, Sandro Meloni, and Yamir Moreno (2017). “A Multilayer perspective for the analysis of urban transportation systems”. In: *Scientific Reports* 7.1, pp. 1–9. ISSN: 20452322. DOI: 10.1038/srep44359. arXiv: 1607.00072. URL: www.nature.com/scientificreports.

Von Landesberger, Tatiana et al. (2016). “MobilityGraphs: Visual Analysis of Mass Mobility Dynamics via Spatio-Temporal Graphs and Clustering”. In: *IEEE Transactions on Visualization and Computer Graphics* 22.1, pp. 11–20. ISSN: 10772626. DOI: 10.1109/TVCG.2015.2468111.

Huang, W et al. (2021). “Urban bus accident analysis: based on a tropos goal risk-accident framework considering learning from incidents process”. In: *Reliability Engineering and System Safety* 216. URL: https://www.sciencedirect.com/science/article/pii/S0951832021004348.

Liao, Feixiong, Theo Arentze, and Harry Timmermans (2010). “Supernetwork Approach for Multimodal and Multiactivity Travel Planning”. In: *Transportation Research Record: Journal of the Transportation Research* 2175.2175, pp. 38–46. DOI: 10.3141/2175-05. URL: https://journals.sagepub.com/doi/abs/10.3141/2175-05?casa_token=uZ44-BQ5hIkAAAAA:r7cPNCifTJZDqKcEuSRBM7gYhI3ORutO02QA3PYExPP6ftve-VWuDhbWLELK_ArKCTE4KvS4A5A4.

Sheffi, Yosef (1985). *Urban transportation networks*. United States of America: PRENTICE-HALL INC. Englewood Cliffs New Jersey 07632. ISBN: 0-13-939729-9. URL: file:///C:/Users/61415/Downloads/UTN_Sheffi.pdf.

Benjamins, M., C. Lindveld, and R. Nes (2001). “MULTIMODAL TRAVEL CHOICE MODELLING: A SUPERNETWORK APPROACH”. In: *undefined*.

Arentze, T. and H. Timmermans (2004). “Multistate supernetwork approach to modelling multi-activity, multimodal trip chains”. In: *International Journal of Geographical Information Science* 18.7, pp. 631–651. ISSN: 1365-8816. DOI: 10.1080/13658810410001701978. URL: http://www.tandfonline.com/doi/abs/10.1080/13658810410001701978.

Arentze, Theo A. and Harry J.P. Timmermans (2009). "A need-based model of multi-day, multi-person activity generation". In: *Transportation Research Part B: Methodological* 43.2, pp. 251–265. ISSN: 01912615. DOI: 10.1016/j.trb.2008.05.007.

Liao, Feixiong, Theo A. Arentze, and Harry J.P. Timmermans (2011). "Constructing personalized transportation networks in multi-state supernetworks: a heuristic approach". In: *International Journal of Geographical Information Science* 25.11, pp. 1885–1903. ISSN: 1365-8816. DOI: 10.1080/13658816.2011.556119. URL: http://www.tandfonline.com/doi/abs/10.1080/13658816.2011.556119.

Liao, Feixiong, Theo Arentze, and Harry Timmermans (2013). "Incorporating space-time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling". In: *Transportation Research Part B: Methodological* 55, pp. 41–58. ISSN: 01912615. DOI: 10.1016/j.trb.2013.05.002.

Li, Zhi Chun, William H.K. Lam, and S. C. Wong (2014). "Bottleneck model revisited: An activity-based perspective". In: *Transportation Research Part B: Methodological* 68, pp. 262–287. ISSN: 01912615. DOI: 10.1016/j.trb.2014.06.013. URL: http://dx.doi.org/10.1016/j.trb.2014.06.013.

Ouyang, L Q et al. (2011). "Network User Equilibrium Model for Scheduling Daily Activity Travel Patterns in Congested Networks". In: *Transportation Research Record* 2254, pp. 131–139. DOI: 10.3141/2254-14.

Lam, William H.K. and Yafeng Yin (2001). "An activity-based time-dependent traffic assignment model". In: *Transportation Research Part B: Methodological* 35.6, pp. 549–574. ISSN: 01912615. DOI: 10.1016/S0191-2615(00)00010-2.

Fu, Xiao and William H K Lam (2013). "A network equilibrium approach for modelling activity-travel pattern scheduling problems in multi-modal transit networks with uncertainty". In: *Transportation*. DOI: 10.1007/s11116-013-9470-9.

Liu, Peng et al. (2015). "Dynamic activity-travel assignment in multi-state supernetworks". In: *Transportation Research Part B: Methodological* 81, pp. 656–671. ISSN: 01912615. DOI: 10.1016/j.trb.2015.08.006.

Reilly, William J. (1931). *The law of retail gravitation*. New York.

Dieter, Karl Heinz (1962). "Distribution of Work Trips in Toronto". In: *Journal of the City Planning Division* 88.1, pp. 9–28. DOI: 10.1061/JCPEAW.0000043. URL: https://ascelibrary.org/doi/abs/10.1061/JCPEAW.0000043.

Bouchard, Richard J and Clyde E Pyers (1965). "Use of Gravity Model for Describing Urban Travel". In: *Highway Research Record* 1.88.

Ben, C, R J Bouchard, and C E Sweet Jr (1965). "An Evaluation of Simplified Procedures for Determining Travel Patterns in a Small Urban Area". In: *Highway Research Record* 1.88.

Evans, Suzanne P. (1973). "A relationship between the gravity model for trip distribution and the transportation problem in linear programming". In: *Transportation Research* 7.1, pp. 39–61. ISSN: 00411647. DOI: 10.1016/0041-1647(73)90005-1.

Low, D E (1972). "New approach to transportation systems modeling". In: *trid.trb.org* 26.3. URL: https://trid.trb.org/view/120004.

Organisation for Economic Co-operation and Development (1974). *Urban traffic models : possibilities for simplification / a report prepared by an OECD Research Group*. OECD. ISBN: 9789264112490.

Robillard, Pierre (1975). "Estimating the O-D matrix from observed link volumes". In: *Transportation Research* 9.2-3, pp. 123–128. ISSN: 00411647. DOI: 10.1016/0041-1647(75)90049-0.

Högberg, Per (1976). "Estimation of parameters in models for traffic prediction: A non-linear regression approach". In: *Transportation Research* 10.4, pp. 263–265. ISSN: 00411647. DOI: 10.1016/0041-1647(76)90059-9.

Willumsen, Luis G (1981). "An entropy maximising model for estimating trip matrices from traffic counts". PhD thesis. the University of Leeds.

Thompson, C. A. et al. (2019). "A cumulative gravity model for inter-urban spatial interaction at different scales". In: *Journal of Transport Geography* 79, p. 102461. ISSN: 09666923. DOI: 10.1016/j.jtrangeo.2019.102461. URL: https://www.sciencedirect.com/science/article/pii/S0966692318304289.

Erlander, S., S. Nguyen, and N. F. Stewart (1979). "On the calibration of the combined distribution-assignment model". In: *Transportation Research Part B* 13.3, pp. 259–267. ISSN: 01912615. DOI: 10.1016/0191-2615(79)90017-1.

Sen, Ashish and Tony E. Smith (1995). *Gravity Models of Spatial Interaction Behavior*. 13th ed. Advances in Spatial and Network Economics. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-642.-79882.-5. DOI: 10.1007/978-3-642-79880-1. URL: http://link.springer.com/10.1007/978-3-642-79880-1.

Hyman, G. M. (1969). "The Calibration of Trip Distribution Models". In: *Environment and Planning A: Economy and Space* 1.1, pp. 105–112. ISSN: 0308-518X. DOI: 10.1068/a010105.

Evans, Andrew W. (1971). "The calibration of trip distribution models with exponential or similar cost functions". In: *Transportation Research* 5.1, pp. 15–38. ISSN: 00411647. DOI: 10.1016/0041-1647(71)90004-9.

Hathaway, PJ (1973). "Some statistical problems associated with trip distribution models". In: *PTRC London*.

Williams, Ian (1976). "A comparison of some calibration techniques for doubly constrained models with an exponential cost function". In: *Transportation Research* 10.2, pp. 91–104. ISSN: 00411647. DOI: 10.1016/0041-1647(76)90045-9.

Daly, Andrew (1982). "Estimating choice models containing attraction variables". In: *Transportation Research Part B* 16.1, pp. 5–15. ISSN: 01912615. DOI: 10.1016/0191-2615(82)90037-6. URL: https://reader.elsevier.com/reader/sd/pii/0191261582900376?token=92B102F4234806A76E72B2D1932B0FDAF93CFCD0E581127A2E908EABFC7E81ED80B73EC0AB985F96A81B1725 originRegion=us-east-1&originCreation=20210430014010.

Shen, Guoqiang and Saniye Gizem Aydin (2014). "Origin–destination missing data estimation for freight transportation planning: a gravity model-based regression approach". In: *Transportation Planning and Technology* 37.6, pp. 505–524. ISSN: 0308-1060. DOI: 10.1080/03081060.2014.927665. URL: http://www.tandfonline.com/doi/abs/10.1080/03081060.2014.927665.

Abdel-Aal, Mounir Mahmoud Moghazy (2014). "Calibrating a trip distribution gravity model stratified by the trip purposes for the city of Alexandria". In: *Alexandria Engineering Journal* 53.3, pp. 677–689. ISSN: 11100168. DOI: 10.1016/j.aej.2014.04.006. URL: http://dx.doi.org/10.1016/j.aej.2014.04.006.

Tamblay, Sebastián et al. (2016). "A zonal inference model based on observed smart-card transactions for Santiago de Chile". In: *Transportation Research Part A: Policy and Practice* 84, pp. 44–54. ISSN: 09658564. DOI: 10.1016/j.tra.2015.10.007.

Murchland, John David (1966). "Some remarks on the gravity model of traffic distribution, and an equivalent maximization formulation". In: *London School of Economics and Political Science, Transport Network Theory Unit*.

Wilson, A. G. (1967). "A statistical theory of spatial distribution models". In: *Transportation Research* 1.3, pp. 253–269. ISSN: 00411647. DOI: 10.1016/0041-1647(67)90035-4.

Wilson, A. G (1969). "The Use of Entropy Maximising Models, in the Theory of Trip Distribution, Mode Split and Route Split". In: *Journal of transport economics and policy* 3.1, pp. 108–126.

Wilson, A G (1971). "A Family of Spatial Interaction Models, and Associated Developments". In: *Environment and Planning A: Economy and Space* 3.1, pp. 1–32. ISSN: 0308-518X. DOI: 10.1068/a030001. URL: http://journals.sagepub.com/doi/10.1068/a030001.

Fang, SC, HSJ Tsao Transportation Science, and undefined 1995 (1995). "Linearly-constrained entropy maximization problem with quadratic cost and its applications to transportation planning problems". In: *pubsonline.informs.org* 29.4, pp. 353–365. DOI: 10.1287/trsc.29.4.353. URL: https://pubsonline.informs.org/doi/abs/10.1287/trsc.29.4.353.

Wang, Dian-Hai, Rong-Han Yao, and Chao Jing (2006). "Entropy Models of Trip Distribution". In: *Journal of Urban Planning and Development* 132.1, pp. 29–35. ISSN: 0733-9488. DOI: 10.1061/(asce)0733-9488(2006)132:1(29).

Xie, Chi, Kara M. Kockelman, and S. Travis Waller (2011). "A maximum entropy-least squares estimator for elastic origin-destination trip matrix estimation". In: *Procedia - Social and Behavioral Sciences*. Vol. 17. Elsevier Ltd, pp. 189–212. DOI: 10.1016/j.sbspro.2011.04.514.

Li, Xiang et al. (2011). "Entropy maximization model for the trip distribution problem with fuzzy and random parameters". In: *Journal of Computational and Applied Mathematics* 235.8, pp. 1906–1913. ISSN: 0377-0427. DOI: 10.1016/J.CAM.2010.09.004.

Tang, J et al. (2018). "Inferring driving trajectories based on probabilistic model from large scale taxi GPS data". In: *Physica A: Statistical Mechanics and its Applications* 506, pp. 566–577. URL: https://www.sciencedirect.com/science/article/pii/S0378437118305016.

Huang, Qinshi et al. (2020). "Spatial-temporal Evolution and Determinants of the Belt and Road Initiative: A Maximum Entropy Gravity Model Approach". In: *Chin. Geogra. SciChinese geographical science* 30.5, pp. 839–854. DOI: 10.1007/s11769-020-1144-x. URL: www.springerlink.com/content/1002-0063.

Kraft, Gerald (1962). *Demand for intercity passenger travel in the Washington-Boston Corridor*. Boston: Systems Analysis, Research Corp., distributed by Clearinghouse for Scientific, and Technical Information, p. 149. URL: https://www.google.com.au/books/edition/Demand_for_Intercity_Passenger_Travel_in/5K0gAQAAMAAJ?hl=en.

Stopher, P R and A H Meyburg (1975). *URBAN TRANSPORTATION MODELING AND PLANNING*. Lexington Books.

Gaudry, Marc J.I. and Michael J. Wills (1977). "Estimating the functional form of travel demand models". In: *Transportation Research* 12.4, pp. 257–289. ISSN: 0041-1647. DOI: 10.1016/0041-1647(78)90068-0.

Carey, Malachy, Chris Hendrickson, and Krishnaswami Siddharthan (1981). "METHOD FOR DIRECT ESTIMATION OF ORIGIN/DESTINATION TRIP MATRICES." In: *Transportation Science* 15.1, pp. 32–49. ISSN: 00411655. DOI: 10.1287/trsc.15.1.32.

Carey, Malachy and Riccardo Revelli (1986). "Constrained Estimation of Direct Demand Functions and Trip Matrices". In: *Transportation Science* 20.3, pp. 143–152. ISSN: 0041-1655. DOI: 10.1287/trsc.20.3.143. URL: http://pubsonline.informs.orghttp//www.informs.org.

Rios, A (2001). "Multimodal freight origin-destination matrix estimation". PhD thesis. Cornell University. URL: https://elibrary.ru/item.asp?id=5228725.

Ríos, Alejandro, Linda K. Nozick, and Mark A. Turnquist (2002). "Value of different categories of information in estimating freight origin-destination tables". In: *Transportation Research Record* 1783, pp. 42–48. ISSN: 03611981. DOI: 10.3141/1783-06.

Kepaptsoglou, Konstantinos, Antony Stathopoulos, and Matthew G. Karlaftis (2017). "Ridership estimation of a new LRT system: Direct demand model approach". In: *Journal of Transport Geography* 58, pp. 146–156. ISSN: 0966-6923. DOI: 10.1016/J.JTRANGEO.2016.12.004.

Yan, Xiang, Xinyu Liu, and Xilei Zhao (2020). "Using machine learning for direct demand modeling of ridesourcing services in Chicago". In: *Journal of Transport Geography* 83, p. 102661. ISSN: 0966-6923. DOI: 10.1016/J.JTRANGEO.2020.102661.

Yang, Hai (1995). "Heuristic algorithms for the bilevel origin-destination matrix estimation problem". In: *Transportation Research Part B* 29.4, pp. 231–242. ISSN: 01912615. DOI: 10.1016/0191-2615(95)00003-V.

Tobin, Roger L. and Terry L. Friesz (1988). "Sensitivity analysis for equilibrium network flow". In: *Transportation Science* 22.4, pp. 242–250. ISSN: 00411655. DOI: 10.1287/trsc.22.4.242. URL: http://pubsonline.informs.org.https//doi.org/10.1287/trsc.22.4.242http://www.informs.org.

Cascetta, Ennio and Francesco Russo (1997). "Calibrating aggregate travel demand models with traffic counts: Estimators and statistical performance". In: *Transportation* 24.3, pp. 271–293. ISSN: 00494488. DOI: 10.1023/A:1004968411792. URL: https://link-springer-com.ezproxy.library.sydney.edu.au/article/10.1023/A:1004968411792.

Nielsen, Otto Anker (1998). "Two New Methods for Estimating Trip Matrices from Traffic Counts". In: *Travel Behaviour Research: Updating the State of Play*. Elsevier, pp. 221–250. DOI: 10.1016/b978-008043360-8/50013-3.

Ashok, K. and M. E. Ben-Akiva (2002). "Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows". In: *Transportation Science* 36.2, pp. 184–198. ISSN: 00411655. DOI: 10.1287/trsc.36.2.184.563. URL: http://pubsonline.informs.org198.https//doi.org/10.1287/trsc.36.2.184.563http://www.informs.org.

Bierlaire, M. and F. Crittin (2004). "An efficient algorithm for real-time estimation and prediction of dynamic OD tables". In: *Operations Research* 52.1, pp. 116–127. ISSN: 0030364X. DOI: 10.

1287/opre.1030.0071. URL: http://pubsonline.informs.org.https//doi.org/10.1287/opre.1030.0071http://www.informs.org.

Zhou, Xuesong and Hani S. Mahmassani (2007). “A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework”. In: *Transportation Research Part B: Methodological* 41.8, pp. 823–840. ISSN: 01912615. DOI: 10.1016/j.trb.2007.02.004.

Cantelmo, Guido et al. (2015). “A Two-Steps Dynamic Demand Estimation Approach Sequentially Adjusting Generations and Distributions”. In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*. Vol. 2015-Octob. Institute of Electrical and Electronics Engineers Inc., pp. 1477–1482. ISBN: 9781467365956. DOI: 10.1109/ITSC.2015.241.

Balakrishna, Ramachandran, Moshe Ben-Akiva, and Haris N Koutsopoulos (2003). “Offline Calibration of Dynamic Traffic Assignment Simultaneous Demand-and-Supply Estimation”. In: *Transportation Research Record: Journal of the Transportation Research Board*, pp. 50–58. DOI: 10.3141/2003-07.

Lu, Lu et al. (2015). “An enhanced SPSA algorithm for the calibration of Dynamic Traffic Assignment models”. In: *Transportation Research Part C: Emerging Technologies* 51, pp. 149–166. ISSN: 0968090X. DOI: 10.1016/j.trc.2014.11.006.

Krishnakumari, Panchamy et al. (2019). “A data driven method for OD matrix estimation”. In: *Transportation Research Part C: Emerging Technologies* 113, pp. 38–56. ISSN: 0968090X. DOI: 10.1016/j.trc.2019.05.014.

Ma, Xiaolei et al. (2019). “Parallel architecture of convolutional bi-directional LSTM neural networks for network-wide metro ridership prediction”. In: *ieeexplore.ieee.org* 20.6, pp. 2278 –2288. DOI: 10.1109/TITS.2018.2867042. URL: https://ieeexplore.ieee.org/abstract/document/8536904/.

Ou, Yuming, Adriana Simona Mihaita, and Fang Chen (2020). “Dynamic Train Demand Estimation and Passenger Assignment”. In: *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*. DOI: 10.1109/ITSC45102.2020.9294606.

Shafiei, Sajjad et al. (2021b). “Integrating data-driven and simulation models to predict traffic state affected by road incidents”. In: *https://doi.org/10.1080/19427867.2021.1916284*. DOI: 10.1080/19427867.2021.1916284. URL: https://www.tandfonline.com/doi/abs/10.1080/19427867.2021.1916284.

Hussain, Etikaf, Ashish Bhaskar, and Edward Chung (2021). “Transit OD matrix estimation using smartcard data: Recent developments and future research challenges”. In: *Transportation Research Part C: Emerging Technologies* 125, p. 103044. ISSN: 0968090X. DOI: 10.1016/j.trc.2021.103044. URL: https://www.sciencedirect.com/science/article/pii/S0968090X21000759.

Tavassoli, Ahmad, Mahmoud Mesbah, and Mark Hickman (2018). “Application of smart card data in validating a large-scale multi-modal transit assignment model”. In: *Public Transport* 10.1, pp. 1–21. ISSN: 16137159. DOI: 10.1007/s12469-017-0171-1.

Yang, X et al. (2017). “Origin-destination estimation using probe vehicle trajectory and link counts”. In: *hindawi.com*. URL: https://www.hindawi.com/journals/jat/2017/4341532/abs/.

Jiang, Tian and Hiroshi Warita (2009). *OD VOLUME FLUCTUATION ANALYSIS FOR TOKYO METROPOLITAN EXPRESSWAY BY ETC DATA*. Tech. rep. the Univ. of Tokyo. URL: https://www.researchgate.net/publication/265001140.

Kim, Jinyoung et al. (2014). "Using electronic toll collection data to understand traffic demand". In: *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations* 18.2, pp. 190–203. ISSN: 15472442. DOI: 10.1080/15472450.2013.806858. URL: https://www.tandfonline.com/action/journalInformation?journalCode=gits20.

Iqbal, Md Shahadat et al. (2014). "Development of origin-destination matrices using mobile phone call data". In: *Transportation Research Part C: Emerging Technologies* 40, pp. 63–74. ISSN: 0968090X. DOI: 10.1016/j.trc.2014.01.002.

Alexander, Lauren et al. (2015). "Origin-destination trips by purpose and time of day inferred from mobile phone data". In: *Transportation Research Part C: Emerging Technologies* 58, pp. 240–250. ISSN: 0968090X. DOI: 10.1016/j.trc.2015.02.018.

Chen, Cynthia et al. (2016). "The promises of big data and small data for travel behavior (aka human mobility) analysis". In: *Transportation Research Part C: Emerging Technologies* 68, pp. 285–299. ISSN: 0968090X. DOI: 10.1016/j.trc.2016.04.005. URL: http://dx.doi.org/10.1016/j.trc.2016.04.005.

Ge, Qian and Daisuke Fukuda (2016). "Updating origin-destination matrices with aggregated data of GPS traces". In: *Transportation Research Part C: Emerging Technologies* 69, pp. 291–312. ISSN: 0968090X. DOI: 10.1016/j.trc.2016.06.002. URL: http://dx.doi.org/10.1016/j.trc.2016.06.002.

Gadziński, Jędrzej (2018). "Perspectives of the use of smartphones in travel behaviour studies: Findings from a literature review and a pilot study". In: *Transportation Research Part C: Emerging Technologies* 88, pp. 74–86. ISSN: 0968090X. DOI: 10.1016/j.trc.2018.01.011.

Abdelgawad, Hossam and Baher Abdulhai (2010). "Managing large-scale multimodal emergency evacuations". In: *Journal of Transportation Safety and Security* 2.2, pp. 122–151. ISSN: 19439970. DOI: 10.1080/19439962.2010.487636. URL: https://www.tandfonline.com/doi/abs/10.1080/19439962.2010.487636.

Cox, Andrew, Fynnwin Prager, and Adam Rose (2011). "Transportation security and the role of resilience: A foundation for operational metrics". In: *Transport Policy* 18.2, pp. 307–317. ISSN: 0967-070X. DOI: 10.1016/J.TRANPOL.2010.09.004.

Jang, Won Seok et al. (2014). "Development of a prototype web GIS-based disaster management system for safe operation of the next generation bimodal tram, South Korea-Focused flooding and snowfall". In: *Sustainability (Switzerland)* 6.4, pp. 1776 –1795. ISSN: 20711050. DOI: 10.3390/su6041776. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84899155644&doi=10.3390%2Fsu6041776&partnerID=40&md5=7bf5f18e9d4a97520fe78a366506cce5.

Cats, Oded and Erik Jenelius (2015). "Planning for the unexpected: The value of reserve capacity for public transport network robustness". In: *Transp Res Part A* 81, pp. 47–61. ISSN: 09658564. DOI: 10.1016/j.tra.2015.02.013.

Neumanna, Andreas (2015). "Why closing an airport may not matter - The impact of the relocation of TXL airport on the bus network of Berlin". In: *Procedia Computer Science*. Ed. by Shakshuki E. Vol. 52. 1. Elsevier B.V., pp. 896 –901. DOI: 10.1016/j.procs.2015.05.160. URL: https:

//www.scopus.com/inward/record.uri?eid=2-s2.0-84939136402&doi=10.1016%2Fj.procs.2015.05.160&partnerID=40&md5=3d9cb492c15c91cd1e803690b5a00e82.

Cats, O, M Yap, and N van Oort (2016). "Exposing the role of exposure: Public transport network risk analysis". In: *Transportation Research Part A: Policy and Practice* 88, pp. 1 –14. ISSN: 09658564. DOI: 10.1016/j.tra.2016.03.015. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84962775957&doi=10.1016%2Fj.tra.2016.03.015&partnerID=40&md5=dc11cdfd36ab9d79f5f6f1f9d4d5c31c.

Challender, Joel (2016). "Storm surge impact to subterranean areas by hurricane sandy, and lessons for Japan's storm surge countermeasures". In: *Journal of Disaster Research* 11.2, pp. 274 –284. ISSN: 18812473. DOI: 10.20965/jdr.2016.p0274. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84973660973&doi=10.20965%2Fjdr.2016.p0274&partnerID=40&md5=8db161adcc00c3a4aee8a237d5347157.

Moylan, Emily, Fletcher Foti, and Alexander Skabardonis (2016). "Observed and simulated traffic impacts from the 2013 Bay Area Rapid Transit strike". In: *Transportation Planning and Technology* 39.2, pp. 162–179. ISSN: 10290354. DOI: 10.1080/03081060.2015.1127539. URL: https://www.tandfonline.com/doi/abs/10.1080/03081060.2015.1127539.

Cottrill, Caitlin et al. (2017). "Tweeting Transit: An examination of social media strategies for transport information management during a large event". In: *Transportation Research Part C: Emerging Technologies* 77, pp. 421 –432. ISSN: 0968090X. DOI: 10.1016/j.trc.2017.02.008. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85013278804&doi=10.1016%2Fj.trc.2017.02.008&partnerID=40&md5=f9e5915f02e5c149c3f2d511ec9ddc4d.

D'Andrea, Eleonora and Francesco Marcelloni (2017). "Detection of traffic congestion and incidents from GPS trace analysis". In: *Expert Systems with Applications* 73, pp. 43–56. ISSN: 0957-4174. DOI: 10.1016/J.ESWA.2016.12.018.

Shelat, Sanmay and Oded Cats (2017). "Measuring spill-over effects of disruptions in public transport networks". In: *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., pp. 756–761. ISBN: 9781509064847. DOI: 10.1109/MTITS.2017.8005613.

Yap, M. D. et al. (2017). "A robust transfer inference algorithm for public transport journeys during disruptions". In: *Transportation Research Procedia*. Vol. 27. Elsevier B.V., pp. 1042–1049. DOI: 10.1016/j.trpro.2017.12.099. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85039996609&doi=10.1016%2Fj.trpro.2017.12.099&partnerID=40&md5=62cf44ff82c392c7cd20c069fbf3865c.

Cats, Oded and Erik Jenelius (2018). "Beyond a complete failure: the impact of partial capacity degradation on public transport network vulnerability". In: *Transportmetrica B* 6.2, pp. 77–96. ISSN: 21680566. DOI: 10.1080/21680566.2016.1267596. URL: https://www.tandfonline.com/action/journalInformation?journalCode=ttrb20https://www.scopus.com/inward/record.uri?eid=2-s2.0-85006124214&doi=10.1080%2F21680566.2016.1267596&partnerID=40&md5=fadcb88f49188f3352a70f32ec237c15https://www.tandfonline.com/doi/.

Diab, Ehab, Guangnan Feng, and Amer Shalaby (2018). "Breaking into emergency shuttle service: aspects and impacts of retracting buses from existing scheduled bus services". In: *Canadian journal of civil engineering* 45.8, pp. 647–658. ISSN: 12086029. DOI: 10.1139/cjce-2017-0294.

Diab, Ehab and Amer Shalaby (2018). "Subway Service Down Again? Assessing the Effects of Subway Service Interruptions on Local Surface Transit Performance". In: *Transportation Research Record* 2672.8, pp. 443–454. ISSN: 21694052. DOI: 10.1177/0361198118791665.

Malandri, Caterina, Achille Fonzone, and Oded Cats (2018). "Recovery time and propagation effects of passenger transport disruptions". In: *Physica A: Statistical Mechanics and its Applications* 505, pp. 7–17. ISSN: 03784371. DOI: 10.1016/j.physa.2018.03.028.

Oswald Beiler, Michelle R, M Asce, and Greg Miller (2018). "Sea-Level Rise Vulnerability Assessment of Bicycle and Trail Networks". In: *Journal of Transportation Engineering, Part A: Systems* 145.2, p. 05018003. ISSN: 2473-2907. DOI: 10.1061/JTEPBS.0000214. URL: https://ascelibrary.org/doi/abs/10.1061/JTEPBS.0000214.

Nguyen Phuoc, Duy Q. et al. (2018). "Exploring the impact of public transport strikes on travel behavior and traffic congestion". In: *International Journal of Sustainable Transportation* 12.8, pp. 613–623. ISSN: 15568334. DOI: 10.1080/15568318.2017.1419322. URL: https://www.tandfonline.com/doi/abs/10.1080/15568318.2017.1419322.

Yang, Saini et al. (2018). "Criticality ranking for components of a transportation network at risk from tropical cyclones". In: *International Journal of Disaster Risk Reduction* 28, pp. 43–55. ISSN: 2212-4209. DOI: 10.1016/J.IJDRR.2018.02.017.

Yap, M. D. et al. (2018). "Identification and quantification of link vulnerability in multi-level public transport networks: a passenger perspective". In: *Transportation* 45.4, pp. 1161–1180. ISSN: 15729435. DOI: 10.1007/S11116-018-9892-5/FIGURES/6. URL: https://link.springer.com/article/10.1007/s11116-018-9892-5https://www.scopus.com/inward/record.uri?eid=2-s2.0-85047830252&doi=10.1007%2Fs11116-018-9892-5&partnerID=40&md5=455a5a8c15eff0cc146f390096d20ae9.

Ari Wibowo, Agus Urip, Selo Sulistyo, and I. Wayan Mustika (2019). "Study on a Train-and Bus-based Delay-Tolerant Networks: Scheduled Mobility and Impact on Routing". In: *Proceeding - 2019 International Conference of Artificial Intelligence and Information Technology, ICAIIT 2019*, pp. 384–390. DOI: 10.1109/ICAIIT.2019.8834675.

Saxena, Neeraj, Taha Hossein Rashidi, and Joshua Auld (2019). "Studying the tastes effecting mode choice behavior of travelers under transit service disruptions". In: *Travel Behaviour and Society* 17, pp. 86–95. ISSN: 2214367X. DOI: 10.1016/j.tbs.2019.07.004.

Azolin, Luiza Gagno, Antônio Nélson da Silva, and Nuno Pinto (2020). "Incorporating public transport in a methodology for assessing resilience in urban mobility". In: *Transportation Research Part D: Transport and Environment* 85. ISSN: 13619209. DOI: 10.1016/j.trd.2020.102386. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85085738030&doi=10.1016%2Fj.trd.2020.102386&partnerID=40&md5=232aaa1d7654b4cbfc9914b678028482.

Consilvio, Alice et al. (2020). "Design of collaborative multimodal strategies for disruption management in urban railway systems". In: *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*. Institute of Electrical and Electronics Engineers Inc., pp. 1–6. ISBN: 9781728141497. DOI: 10.1109/ITSC45102.2020.9294329.

Fang, Yan, Yiping Jiang, and Wei Fei (2020). "Disruption Recovery for Urban Public Tram System: An Analysis of Replacement Service Selection". In: *IEEE Access* 8, pp. 31633–31646. ISSN: 21693536. DOI: 10.1109/ACCESS.2020.2972445.

Jia, Jianmin, Hui Zhang, and Baiying Shi (2020). "Exploring Bike-Sharing Behavior Affected by Public Transportation Disruption: Case of Washington, DC, Metro Shutdown". In: *Journal of Transportation Engineering, Part A: Systems* 147.3, p. 04020163. ISSN: 2473-2907. DOI: 10.1061/JTEPBS.0000495. URL: https://ascelibrary.org/doi/abs/10.1061/JTEPBS.0000495.

Leng, Nuannuan and Francesco Corman (2020b). "The role of information availability to passengers in public transport disruptions: An agent-based simulation approach". In: *Transportation Research Part A: Policy and Practice* 133, pp. 214 –236. ISSN: 09658564. DOI: 10.1016/j.tra.2020.01.007. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85078999823&doi=10.1016%2Fj.tra.2020.01.007&partnerID=40&md5=7deade161ec3857bccc88f831e9be960.

— (2020a). "How the issue time of information affects passengers in public transport disruptions: An agent-based simulation approach". In: *Procedia Computer Science*. Ed. by Shakshuki E. and Yasar A.-U.-H. Vol. 170. Elsevier B.V., pp. 382 –389. DOI: 10.1016/j.procs.2020.03.068. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85085578829&doi=10.1016%2Fj.procs.2020.03.068&partnerID=40&md5=f766da00149e05f6b3f209dca2743053.

Marra, Alessio Daniele and Francesco Corman (2020). "From Delay to Disruption: Impact of Service Degradation on Public Transport Networks". In: *Transportation Research Record* 2674.10, pp. 886 –897. ISSN: 03611981. DOI: 10.1177/0361198120940989. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85102650748&doi=10.1177%2F0361198120940989&partnerID=40&md5=91878cad9708ebe5b4d30df46e80a68b.

Rahimi, Mahsa and Francesco Corman (2020). "Quantifying the heterogeneity of direct and indirect effects of a public transport disruption with agent-based simulation". In: *20th Swiss Transport Research Conference*. STRC, pp. 1–42. DOI: 10.3929/ETHZ-B-000419839.

Tan, Zhijia et al. (2020). "Evacuating metro passengers via the urban bus system under uncertain disruption recovery time and heterogeneous risk-taking behaviour". In: *Transportation Research Part C: Emerging Technologies* 119. DOI: 10.1016/J.TRC.2020.102761.

Aparicio, Joao T, Elisabete Arsenio, and Rui Henriques (2021). "Understanding the impacts of the COVID-19 pandemic on public transportation travel patterns in the city of Lisbon". In: *Sustainability (Switzerland)* 13.15. ISSN: 20711050. DOI: 10.3390/su13158342. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85111665710&doi=10.3390%2Fsu13158342&partnerID=40&md5=946da6adf5b26a41d197eb648750f8ca.

Barbieri, Diego Maria et al. (2021). "Impact of COVID-19 pandemic on mobility in ten countries and associated perceived risk for all transport modes". In: *PLoS ONE* 16.2 February. ISSN: 19326203. DOI: 10.1371/journal.pone.0245886. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85100549136&doi=10.1371%2Fjournal.pone.0245886&partnerID=40&md5=4b7d9cdcfc9562d0e953c3ca577ef57e.

Cebecauer, Matej et al. (2021). "Integrating Demand Responsive Services into Public Transport Disruption Management". In: *IEEE Open Journal of Intelligent Transportation Systems* 2, pp. 24–36. ISSN: 26877813. DOI: 10.1109/OJITS.2021.3057221. URL: https://www.scopus.com/

inward/record.uri?eid=2-s2.0-85141361403&doi=10.1109%2FOJITS.2021.3057221&partnerID=40&md5=df7841fbbdcc36015525a937ce416d47.

Drabicki, Arkadiusz Adam, Md Faqhrul Islam, and Andrzej Szarata (2021). "Investigating the impact of public transport service disruptions upon passenger travel behaviour—Results from Krakow City". In: *Energies* 14.16. ISSN: 19961073. DOI: 10.3390/en14164889. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85112722813&doi=10.3390%2Fen14164889&partnerID=40&md5=fcef50bb70fb57e3f4376f5723068aec.

He, Yiyi et al. (2021). "Flood impacts on urban transit and accessibility—A case study of Kinshasa". In: *Transportation Research Part D: Transport and Environment* 96, p. 102889. ISSN: 13619209. DOI: 10.1016/j.trd.2021.102889. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85107023837&doi=10.1016%2Fj.trd.2021.102889&partnerID=40&md5=672809e451365bd480473d7543d93c36.

Henry, Elise, Angelo Furno, and Nour-Eddin El Faouzi (2021). "REINFORCE: rapid augmentation of large-scale multi-modal transport networks for resilience enhancement". In: *Applied Network Science* 6.1. ISSN: 23648228. DOI: 10.1007/s41109-021-00422-2. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85117699467&doi=10.1007%2Fs41109-021-00422-2&partnerID=40&md5=f3059f08c7f2a3d57770942717d6e89f.

Rahimi, Mahsa ; et al. (2021). "Multi-modal management actions for public transport disruptions: an agent-based simulation". In: *ETH*. ETH, pp. 1–7. ISBN: 978-172818995-6. DOI: 10.3929/ethz-b-000501419. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85115825930&doi=10.1109%2FMT-ITS49943.2021.9529305&partnerID=40&md5=f9d76aa43788651300e235bf85d5264ahttps://doi.org/10.3929/ethz-b-000501419.

Cong, Chengli et al. (2022). "Impact Estimation of Unplanned Urban Rail Disruptions on Public Transport Passengers: A Multi-Agent Based Simulation Approach". In: *International Journal of Environmental Research and Public Health* 19.15. ISSN: 16604601. DOI: 10.3390/ijerph19159052. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135379063&doi=10.3390%2Fijerph19159052&partnerID=40&md5=2903d1eb406f730e62f6552d495876af.

Mo, Baichuan, Haris N. Koutsopoulos, and Jinhua Zhao (2022). "Inferring passenger responses to urban rail disruptions using smart card data: A probabilistic framework". In: *Transportation Research Part E: Logistics and Transportation Review* 159, p. 102628. ISSN: 13665545. DOI: 10.1016/j.tre.2022.102628.

Yang, Yuanxuan et al. (2022). "Understanding the impacts of public transit disruptions on bikeshare schemes and cycling behaviours using spatiotemporal and graph-based analysis: A case study of four London Tube strikes". In: *Journal of Transport Geography* 98. ISSN: 09666923. DOI: 10.1016/j.jtrangeo.2021.103255. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85121916800&doi=10.1016%2Fj.jtrangeo.2021.103255&partnerID=40&md5=a0b0bb1b574d09b9c4f497d21759a35f.

Yap, Menno and Oded Cats (2022). "Analysis and prediction of ridership impacts during planned public transport disruptions". In: *Journal of Public Transportation* 24. ISSN: 1077291X. DOI: 10.1016/j.jpubtr.2022.100036. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85138546287&doi=10.1016%2Fj.jpubtr.2022.100036&partnerID=40&md5=117878a7f459f7d7555a79104c884b31.

Yap, Menno et al. (2022). “Quantification and control of disruption propagation in multi-level public transport networks”. In: *International Journal of Transportation Science and Technology* 11.1, pp. 83–106. ISSN: 20460449. DOI: 10.1016/j.ijtst.2021.02.002. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85105259516&doi=10.1016%2Fj.ijtst.2021.02.002&partnerID=40&md5=7359e29feb139d3d101ce6811de76777.

Böcker, Lars et al. (2023). “Pandemic impacts on public transport safety and stress perceptions in Nordic cities”. In: *Transportation Research Part D: Transport and Environment* 114. ISSN: 13619209. DOI: 10.1016/j.trd.2022.103562. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85144491384&doi=10.1016%2Fj.trd.2022.103562&partnerID=40&md5=5a41c114b6cf6fecc5eef798b7144a53.

Drabicki, Arkadiusz et al. (2023). “Should I stay or should I board? Willingness to wait with real-time crowding information in urban public transport”. In: *Research in Transportation Business and Management* 47. ISSN: 22105395. DOI: 10.1016/j.rtbm.2023.100963. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85150413684&doi=10.1016%2Fj.rtbm.2023.100963&partnerID=40&md5=5e7870e75507f5e16ab0318f9cfaee4d.

Feng, Yinshuai et al. (2023). “Multi-scale risk assessment and mitigations comparison for COVID-19 in urban public transport: A combined field measurement and modeling approach”. In: *Building and Environment* 242. ISSN: 03601323. DOI: 10.1016/j.buildenv.2023.110489. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85162042280&doi=10.1016%2Fj.buildenv.2023.110489&partnerID=40&md5=75683aa5c2290cb6079c2a2105654063.

Hasiak, Sophie (2023). “Public Transport Mobility and COVID-19 Health Crisis: What Changes in Students’ Mobility Behaviors?” In: *Transportation Research Procedia*. Ed. by de Picado Santos L., Pinho de Sousa J., and Arsenio E. Vol. 72. Elsevier B.V., pp. 1531–1538. DOI: 10.1016/j.trpro.2023.11.620. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85182945173&doi=10.1016%2Fj.trpro.2023.11.620&partnerID=40&md5=282b7f91360680416d1299c2b469d7e5https://linkinghub.elsevier.com/retrieve/pii/S2352146523009183.

Marra, Alessio Daniele and Francesco Corman (2023). “How different network disturbances affect route choice of public transport passengers. A descriptive study based on tracking”. In: *Expert Systems with Applications* 213. ISSN: 09574174. DOI: 10.1016/j.eswa.2022.119083. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85140648359&doi=10.1016%2Fj.eswa.2022.119083&partnerID=40&md5=eacccad6067121d76eb2bfa41be3256f.

Othman, Muhammad Shalihin et al. (2023). “Real-Time Simulation Framework with Traffic Incident Prediction: A Singapore Case Study”. In: *Proceedings - 2023 IEEE/ACM 27th International Symposium on Distributed Simulation and Real Time Applications, DS-RT 2023*. Institute of Electrical and Electronics Engineers Inc., pp. 84–90. ISBN: 9798350337846. DOI: 10.1109/DS-RT58998.2023.00019. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85179135918&doi=10.1109%2FDS-RT58998.2023.00019&partnerID=40&md5=f7a68a92e92c2f279e0e7c6055e7bad0.

Soza-Parra, Jaime, Ignacio Tiznado-Aitken, and Juan Carlos Muñoz (2023). “A discrete-event public transportation simulation model to evaluate travel demand management impacts on waiting times and crowding conditions”. In: *Journal of Public Transportation* 25. ISSN: 1077291X. DOI: 10.

1016/j.jpubtr.2023.100075. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85175421171&doi=10.1016%2Fj.jpubtr.2023.100075&partnerID=40&md5=e159389c6e7dc96ad99a5d618745a0bc.

He, Sylvia Y., Sui Tao, and Ka Kit Sun (2024). "Attitudes towards public transport under extended disruptions and massive-scale transit dysfunction: A Hong Kong case study". In: *Transport Policy* 149, pp. 247–258. ISSN: 1879310X. DOI: 10.1016/j.tranpol.2024.02.008. URL: https://www-sciencedirect-com.ezproxy.lib.uts.edu.au/science/article/pii/S0967070X24000441?via%3Dihub.

Zhang, Li et al. (2024a). "Analysis of multi-modal public transportation system performance under metro disruptions: A dynamic resilience assessment framework". In: *Transportation Research Part A: Policy and Practice* 183, p. 104077. ISSN: 09658564. DOI: 10.1016/j.tra.2024.104077. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85190589542&doi=10.1016%2Fj.tra.2024.104077&partnerID=40&md5=37585d0599e7fb90bf9ff30d4a4d4037.

Zhu, Weihua et al. (2022). "System vulnerability to flood events and risk assessment of railway systems based on national and river basin scales in China". In: *Natural Hazards and Earth System Sciences* 22.5, pp. 1519 –1540. ISSN: 15618633. DOI: 10.5194/nhess-22-1519-2022. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85130180031&doi=10.5194%2Fnhess-22-1519-2022&partnerID=40&md5=c36035adeaf71660f1f126d3191b6def.

Zhu, Yuan et al. (2019). "Modeling of Incident-Induced Capacity Loss for Hurricane Evacuation Simulation". In: *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*. Institute of Electrical and Electronics Engineers Inc., pp. 613–618. ISBN: 9781538670248. DOI: 10.1109/ITSC.2019.8917250.

Hammond, M. J. et al. (2015). "Urban flood impact assessment: A state-of-the-art review". In: *Urban Water Journal* 12.1, pp. 14–29. DOI: 10.1080/1573062X.2013.857421.

Munawar, Hafiz Suliman, Ahmed W.A. Hammad, and S. Travis Waller (2021). "A review on flood management technologies related to image processing and machine learning". In: *Automation in Construction* 132, p. 103916. ISSN: 0926-5805. DOI: 10.1016/J.AUTCON.2021.103916.

Bhuiyan, Tariqur Rahman et al. (2021). "The socioeconomic impact of climate-related hazards: flash flood impact assessment in Kuala Lumpur, Malaysia". In: *Natural Hazards* 109.2, pp. 1509–1538. ISSN: 15730840. DOI: 10.1007/s11069-021-04887-3. URL: https://link.springer.com/article/10.1007/s11069-021-04887-3.

Pregnolato, Maria et al. (2017). "The impact of flooding on road transport: A depth-disruption function". In: *Transportation Research Part D Transport and Environment* 55, pp. 67–81. ISSN: 13619209. DOI: 10.1016/j.trd.2017.06.020.

Ansari, Abdullah, K. Seshagiri Rao, and Arvind K. Jain (2024). "Application of Microzonation Towards System-Wide Seismic Risk Assessment of Railway Network". In: *Transportation Infrastructure Geotechnology* 11.3, pp. 1119–1142. ISSN: 21967210. DOI: 10.1007/s40515-023-00317-y. URL: https://doi.org/10.1007/s40515-023-00317-y.

Ronchi, Enrico et al. (2019). "An open multi-physics framework for modelling wildland-urban interface fire evacuations". In: *Safety Science* 118, pp. 868–880. ISSN: 18791042. DOI: 10.1016/j.ssci.2019.06.009.

Wahlqvist, Jonathan et al. (2021). "The simulation of wildland-urban interface fire evacuation: The WUI-NITY platform". In: *Safety Science* 136, p. 105145. ISSN: 0925-7535. DOI: 10.1016/J.SSCI.2020.105145.

Cova, Thomas J. et al. (2021). "Toward Simulating Dire Wildfire Scenarios". In: *Natural Hazards Review* 22.3, p. 06021003. ISSN: 1527-6988. DOI: 10.1061/(ASCE)NH.1527-6996.0000474. URL: https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29NH.1527-6996.0000474.

Lovreglio, Ruggiero et al. (2020). "Calibrating the Wildfire Decision Model using hybrid choice modelling". In: *International Journal of Disaster Risk Reduction* 50, p. 101770. ISSN: 2212-4209. DOI: 10.1016/J.IJDRR.2020.101770.

Ren, Rui et al. (2019). "Statistical analysis of fire accidents in Chinese highway tunnels 2000–2016". In: *Tunnelling and Underground Space Technology* 83, pp. 452–460. ISSN: 0886-7798. DOI: 10.1016/J.TUST.2018.10.008.

Bjelland, Henrik et al. (2021). "Emergency preparedness for tunnel fires – A systems-oriented approach". In: *Safety Science* 143, p. 105408. ISSN: 0925-7535. DOI: 10.1016/J.SSCI.2021.105408.

Wetterberg, Niklas, Enrico Ronchi, and Jonathan Wahlqvist (2021). "Individual Driving Behaviour in Wildfire Smoke". In: *Fire Technology* 57.3, pp. 1041–1061. ISSN: 15728099. DOI: 10.1007/S10694-020-01026-5/FIGURES/7. URL: https://link.springer.com/article/10.1007/s10694-020-01026-5.

Yasanthi, Navoda and Rillagodage Regina (2018). "Investigating the Impacts of Adverse Road Weather Conditions on Free-Flow Speed and Lane Utilization". PhD thesis. University of Regina.

Inatsu, Masaru, Seika Tanji, and Yousuke Sato (2020). "Toward predicting expressway closures due to blowing snow events". In: *Cold Regions Science and Technology* 177, p. 103123. ISSN: 0165-232X. DOI: 10.1016/J.COLDREGIONS.2020.103123.

Elhenawy, Mohammed, Hesham A. Rakha, and Huthaifa I. Ashqar (2021). "Joint Impact of Rain and Incidents on Traffic Stream Speeds". In: *Journal of Advanced Transportation* 2021. ISSN: 20423195. DOI: 10.1155/2021/8812740.

Page-Tan, Courtney and Daniel P Aldrich (2024). "Snowed in? An analysis of social and transportation networks in the 2015 Boston Snowmageddon". In: *Sustainable Development* 32.2, pp. 1446 – 1457. ISSN: 09680802. DOI: 10.1002/sd.2680. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85165495578&doi=10.1002%2Fsd.2680&partnerID=40&md5=6e1a4790f7b8af9802430be5e54371a4.

Jarmuz, Dominika and Jerzy Chmiel (2020). "A review of approaches to the study of Weather's effect on road traffic parameters". In: *Transport Problems* 15.4, pp. 241–251. ISSN: 2300861X. DOI: 10.21307/TP-2020-063. URL: https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-8272f8cc-6d33-40c8-b645-5f41336e394b.

Garcia-Arteaga, Juan D and Laura Lotero (2023). "A network-based analysis to assess COVID-19 disruptions in the Bogotá BRT system". In: *Environment and Planning B: Urban Analytics and City Science* 50.4, pp. 983 –999. ISSN: 23998083. DOI: 10.1177/23998083221150646. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85146610080&doi=10.1177%2F23998083221150646&partnerID=40&md5=67085520e29ca0e49a8efbf31f80705b.

Du, Bo et al. (2024). "Impact of the COVID-19 pandemic on daily travel: Findings from New South Wales, Australia". In: *Travel Behaviour and Society* 36. ISSN: 2214367X. DOI: 10.1016/j.tbs.2024.100798. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85189519313&doi=10.1016%2Fj.tbs.2024.100798&partnerID=40&md5=1deabd71c0a2c9316e8f2e820a99e1d2.

Li, Ruimin, Francisco C. Pereira, and Moshe E. Ben-Akiva (2018). "Overview of traffic incident duration analysis and prediction". In: *European Transport Research Review* 10.2, pp. 1–13. ISSN: 18668887. DOI: 10.1186/S12544-018-0300-1/TABLES/4. URL: https://etrr.springeropen.com/articles/10.1186/s12544-018-0300-1.

Yang, Zekun, Wenping Zhang, and Juan Feng (2022). "Predicting multiple types of traffic accident severity with explanations: A multi-task deep learning framework". In: *Safety Science* 146, p. 105522. ISSN: 0925-7535. DOI: 10.1016/J.SSCI.2021.105522.

Yap, Menno and Oded Cats (2021). "Predicting disruptions and their passenger delay impacts for public transport stops". In: *Transportation* 48.4, pp. 1703–1731. ISSN: 00494488. DOI: 10.1007/S11116-020-10109-9. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85084236778&doi=10.1007%2Fs11116-020-10109-9&partnerID=40&md5=486d5ed5a546337aacf523896488518c.

Wen, Tao et al. (2018). "Integrated Incident Decision-Support using Traffic Simulation and Data-Driven Models". In: *Transportation Research Record* 2672.42, pp. 247–256. ISSN: 0361-1981. DOI: 10.1177/0361198118782270. URL: http://journals.sagepub.com/doi/10.1177/0361198118782270.

Liu, Kai, Jiatong Zhu, and Ming Wang (2021). "An event-based probabilistic model of disruption risk to urban metro networks". In: *Transportation Research Part A: Policy and Practice* 147, pp. 93–105. ISSN: 09658564. DOI: 10.1016/j.tra.2021.03.010.

Zhao, Yangyang et al. (2021). "Short-Term Metro Ridership Prediction During Unplanned Events:" in: *Transportation Research Record*, p. 036119812110375. ISSN: 0361-1981. DOI: 10.1177/03611981211037553. URL: https://journals.sagepub.com/doi/full/10.1177/03611981211037553?casa_token=B_J6YIa55wgAAAAA%3AmVcPXMguUbQ_DP0ge7c4FFgcFtlgeht_yjIDOepJh_lzvpZsFpusp2GFfqK3VJJi4iGGhyiB3hI.

Zhao, Dong et al. (2023). "Modelling public transport disruptions and impact by smart-card data". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*. September, pp. 2945–2952. ISBN: 9798350399462. DOI: 10.1109/ITSC57777.2023.10422238.

Eltved, Morten et al. (2021). "Impacts of long-term service disruptions on passenger travel behaviour: A smart card analysis from the Greater Copenhagen area". In: *Transportation Research Part C: Emerging Technologies* 131. ISSN: 0968090X. DOI: 10.1016/j.trc.2021.103198. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85112465721&doi=10.1016%2Fj.trc.2021.103198&partnerID=40&md5=fec92229bdad12c8a1720f4463d3e39a.

Shafiei, Sajjad et al. (2021a). "Arterial corridor travel time prediction under non-recurring conditions". In: *Journal of Intelligent Transportation Systems*. ISSN: 15472442. DOI: 10.1080/15472450.2021.2023017. URL: https://www.tandfonline.com/doi/abs/10.1080/15472450.2021.2023017.

Müller, Sebastian Alexander, Gregor Leich, and Kai Nagel (2020). "The effect of unexpected disruptions and information times on public transport passengers: A simulation study". In: *Procedia Computer Science*. Ed. by Shakshuki E. and Yasar A.-U.-H. Vol. 170. Elsevier B.V., pp. 745 –750. DOI: 10.1016/j.procs.2020.03.161. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85085576821&doi=10.1016%2Fj.procs.2020.03.161&partnerID=40&md5=f84da6ccd73349d710c5ff43b968aeeb.

Lu, Qing Chang (2018). "Modeling network resilience of rail transit under operational incidents". In: *Transportation Research Part A: Policy and Practice* 117, pp. 227–237. ISSN: 0965-8564. DOI: 10.1016/J.TRA.2018.08.015.

Lu, Jian, Xiaoli Ma, and Yingying Xing (2021). "Risk factors affecting the severity of disruptions in metro operation in shanghai, 2013-2016". In: *Journal of Transportation Safety and Security* 13.1, pp. 69–92. DOI: 10.1080/19439962.2019.1605640.

Yang, Zhen, Xiaobin Dong, and Li Guo (2023). "Scenario inference model of urban metro system cascading failure under extreme rainfall conditions". In: *Reliability Engineering and System Safety* 229. ISSN: 09518320. DOI: 10.1016/j.ress.2022.108888.

Borowski, Elisa et al. (2023). "Does ridesourcing respond to unplanned rail disruptions? A natural experiment analysis of mobility resilience and disparity". In: *Cities* 140. ISSN: 02642751. DOI: 10.1016/j.cities.2023.104439. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85163737753&doi=10.1016%2Fj.cities.2023.104439&partnerID=40&md5=db24e6f35674b93fc5f877728fc29ba7.

Papangelis, Konstantinos et al. (2013). "Developing a real time passenger information system for rural areas". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8017 LNCS.PART 2, pp. 153 –162. ISSN: 16113349. DOI: 10.1007/978-3-642-39215-3_19. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84880863972&doi=10.1007%2F978-3-642-39215-3_19&partnerID=40&md5=41fc65059f1d8a02a8eb0246e6b50e3e.

Sun, Huijun et al. (2016). "Estimating the influence of common disruptions on urban rail transit networks". In: *Transportation Research Part A: Policy and Practice* 94, pp. 62–75. ISSN: 09658564. DOI: 10.1016/j.tra.2016.09.006.

Liu, Tianyou, Zhenliang Ma, and Haris N. Koutsopoulos (2021). "Unplanned Disruption Analysis in Urban Railway Systems Using Smart Card Data". In: *Urban Rail Transit* 7.3, pp. 177–190. ISSN: 21996679. DOI: 10.1007/s40864-021-00150-x. URL: https://link.springer.com/article/10.1007/s40864-021-00150-x.

Yap, Menno and Oded Cats (2019). "Analysis and Prediction of Disruptions in Metro Networks". In: *MT-ITS 2019 - 6th International Conference on Models and Technologies for Intelligent Transportation Systems*. Institute of Electrical and Electronics Engineers Inc., p. 1. ISBN: 978-153869484-8. DOI: 10.1109/MTITS.2019.8883320. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85074947989&doi=10.1109%2FMTITS.2019.8883320&partnerID=40&md5=9c22ed4fbedea92c8089f091c43c27c7.

Zhu, Shanjiang et al. (2017). "Travel behavior reactions to transit service disruptions: Study of metro safetrack projects in Washington, D.C." In: *Transportation Research Record* 2649.1, pp. 79–88. ISSN: 21694052. DOI: 10.3141/2649-09.

Khadilkar, Harshad (2017). "Data-Enabled Stochastic Modeling for Evaluating Schedule Robustness of Railway Networks". In: *Transportation Science* 51.4, pp. 1161–1176. ISSN: 1526-5447. DOI: 10.1287/trsc.2016.0703. URL: http://pubsonline.informs.org.https//doi.org/10.1287/trsc.2016.0703http://www.informs.org.

Mahdavi, S M Hassan M, Neila Bhouri, and Gerard Scemama (2020). "Dynamic Resilience of Public Transport Network: A Case Study for Fleet-Failure in Bus Transport Operation of New Delhi". In: *Transportation Research Procedia*. Ed. by Codina E. et al. Vol. 47. Elsevier B.V., pp. 672 –679. DOI: 10.1016/j.trpro.2020.03.146. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85084635027&doi=10.1016%2Fj.trpro.2020.03.146&partnerID=40&md5=b2605a2b59cae2010a6cdc2b49542363.

Zhang, Nan et al. (2022). "Detecting metro service disruptions via large-scale vehicle location data". In: *Transportation Research Part C: Emerging Technologies* 144, p. 103880. ISSN: 0968-090X. DOI: 10.1016/J.TRC.2022.103880.

Tokuda, Yasuharu et al. (2006). "Prehospital management of sarin nerve gas terrorism in urban settings: 10 Years of progress after the Tokyo subway sarin attack". In: *Resuscitation* 68.2, pp. 193–202. ISSN: 03009572. DOI: 10.1016/j.resuscitation.2005.05.023.

Wang, Yihui et al. (2021b). "Real-time integrated train rescheduling and rolling stock circulation planning for a metro line under disruptions". In: *Transportation Research Part B: Methodological* 152, pp. 87–117. ISSN: 0191-2615. DOI: 10.1016/J.TRB.2021.08.003.

Liang, Jinpeng et al. (2019). "Robust bus bridging service design under rail transit system disruptions". In: *Transportation Research Part E: Logistics and Transportation Review* 132, pp. 97–116. ISSN: 1366-5545. DOI: 10.1016/J.TRE.2019.10.008.

Kepaptsoglou, Konstantinos and Matthew G. Karlaftis (2009). "The bus bridging problem in metro operations: conceptual framework, models and algorithms". In: *Public Transport* 1.4, pp. 275–297. ISSN: 1866749X. DOI: 10.1007/s12469-010-0017-6. URL: https://link.springer.com/content/pdf/10.1007/s12469-010-0017-6.pdf.

Jin, Jian Gang, Kwong Meng Teo, and Amedeo R. Odoni (2016). "Optimizing bus bridging services in response to disruptions of urban transit rail networks". In: *Transportation Science* 50.3, pp. 790–804. ISSN: 15265447. DOI: 10.1287/trsc.2014.0577. URL: https://pubsonline.informs.org/doi/abs/10.1287/trsc.2014.0577.

Zhu, Yongqiu and Rob M.P. Goverde (2019). "Dynamic Passenger Assignment for Major Railway Disruptions Considering Information Interventions". In: *Networks and Spatial Economics* 19.4, pp. 1249–1279. ISSN: 15729427. DOI: 10.1007/s11067-019-09467-4.

Lu, Qing-Chang and Shan Lin (2019). "Vulnerability analysis of urban rail transit network within multi-modal public transport networks". In: *Sustainability (Switzerland)* 11.7. ISSN: 20711050. DOI: 10.3390/SU11072109. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85068027337&doi=10.3390%2FSU11072109&partnerID=40&md5=25540f7ed035bae69ae7a213985d2

Xu, Mengtong and Xiao Liang (2020). "An optimization bus dispatching model under emergency of urban rail transit system". In: *Proceedings - 2020 International Conference on Urban Engineering and Management Science, ICUEMS 2020*. Institute of Electrical and Electronics Engineers Inc., pp. 390–396. ISBN: 9781728188324. DOI: 10.1109/ICUEMS50872.2020.00089.

Leng, Nuannuan, Zhengwen Liao, and Francesco Corman (2020). "Role of timetable, rolling stock rescheduling, and information strategies to passengers in public transport disruptions". In: *Transportation Research Record* 2674.9, pp. 135–147. ISSN: 21694052. DOI: 10.1177/0361198120927000. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85092295764&doi=10.1177%2F0361198120927000&partnerID=40&md5=19fff099fadf478621d3c78d6e3eb5b4.

Tang, Junqing et al. (2021b). "Multi-disruption resilience assessment of rail transit systems with optimized commuter flows". In: *Reliability Engineering and System Safety* 214, p. 107715. ISSN: 0951-8320. DOI: 10.1016/J.RESS.2021.107715.

Malandri, Caterina et al. (2021). "Public transport network vulnerability and delay distribution among travelers". In: *Sustainability (Switzerland)* 13.16. ISSN: 20711050. DOI: 10.3390/su13168737. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85112213814&doi=10.3390%2Fsu13168737&partnerID=40&md5=08b1053d7f8575c3c040ddf8fccdfd61.

Shao, Lihua and Liying Song (2022). "An Optimization Approach considering Passengers' Space-Time Requirements for Bus Bridging Service under URT Disruption". In: *Journal of Advanced Transportation* 2022. ISSN: 20423195. DOI: 10.1155/2022/2113311.

Tan, Erlong et al. (2023). "Enhancing urban metro system resilience under disruptive events through multi-agent reinforcement learning". In: *Journal of Transportation Safety and Security*. ISSN: 19439970. DOI: 10.1080/19439962.2023.2300272.

Massobrio, Renzo and Oded Cats (2024). "Topological assessment of recoverability in public transport networks". In: *Communications Physics* 7.1. ISSN: 23993650. DOI: 10.1038/s42005-024-01596-8. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85188577025&doi=10.1038%2Fs42005-024-01596-8&partnerID=40&md5=dc36e4664e2b77e9882b9c97f370

Chen, Jinqu et al. (2024). "Urban rail transit disruption management based on passenger guidance and extended bus bridging service considering uncertain bus running time". In: *Expert Systems with Applications* 249. ISSN: 09574174. DOI: 10.1016/j.eswa.2024.123659.

Zhang, Shuang et al. (2024b). "A general metro timetable rescheduling approach for the minimisation of the capacity loss after random line disruption". In: *Transportmetrica A: Transport Science* 20.3. ISSN: 23249943. DOI: 10.1080/23249935.2023.2204965.

Zhang, Yanan et al. (2023). "Vulnerability Analysis of Bus Network Based on Land-Use Type of Bus Stops: The Case of Xi'an, China". In: *Sustainability (Switzerland)* 15.16. ISSN: 20711050. DOI: 10.3390/su151612566. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85169130723&doi=10.3390%2Fsu151612566&partnerID=40&md5=41c99a0c49c0f7911d9023971fb96cb7.

Shi, Jiangang et al. (2019). "Sustainable development of urban rail transit networks: A vulnerability perspective". In: *Sustainability (Switzerland)* 11.5. ISSN: 20711050. DOI: 10.3390/su11051335. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062858353&doi=10.3390%2Fsu11051335&partnerID=40&md5=654de2afea2bc9dbc71d79a7318ab006.

Xing, Yingying et al. (2017). "Vulnerability analysis of urban rail transit based on complex network theory: a case study of Shanghai Metro". In: *Public Transport* 9.3, pp. 501–525. ISSN: 16137159. DOI: 10.1007/S12469-017-0170-2/FIGURES/7. URL: https://link.springer.com/article/10.1007/s12469-017-0170-2.

Zhang, Jianhua, Shuliang Wang, and Xiaoyuan Wang (2018). “Comparison analysis on vulnerability of metro networks based on complex network”. In: *Physica A: Statistical Mechanics and its Applications* 496. ISSN: 03784371. DOI: 10.1016/j.physa.2017.12.094.

Yin, Haodong et al. (2018). “Optimal Bus-Bridging Service under a Metro Station Disruption”. In: *Journal of Advanced Transportation* 2018. DOI: 10.1155/2018/2758652.

Zhang, Dong ming et al. (2018). “Resiliency assessment of urban rail transit networks: Shanghai metro as an example”. In: *Safety Science* 106, pp. 230–243. ISSN: 18791042. DOI: 10.1016/j.ssci.2018.03.023.

Wang, Jiadong, Zhenzhou Yuan, and Yonghao Yin (2019). “Optimization of Bus Bridging Service under Unexpected Metro Disruptions with Dynamic Passenger Flows”. In: *Journal of Advanced Transportation* 2019. DOI: 10.1155/2019/6965728.

Srikukenthiran, Siva and Amer Shalaby (2017). “Enabling large-scale transit microsimulation for disruption response support using the Nexus platform”. In: *Public Transport* 9.1, pp. 411–435. ISSN: 16137159. DOI: 10.1007/s12469-017-0158-y.

Placido, A. P. (2017). “Managing disruptions and disturbances on railway services: a real-scale case study”. In: *International Journal of Transport Development and Integration*, 1(4), 695–710.

Dollevoet, Twan et al. (2018). “Delay Propagation and Delay Management in Transportation Networks”. In: *International Series in Operations Research and Management Science* 268, pp. 285–317. ISSN: 08848289. DOI: 10.1007/978-3-319-72153-8_13.

Dekker, Mark M. and Debabrata Panja (2021). “Cascading dominates large-scale disruptions in transport over complex networks”. In: *PLoS ONE* 16.1 1. DOI: 10.1371/JOURNAL.PONE.0246077.

Corman, Francesco (2020). “Interactions and equilibrium between rescheduling train traffic and routing passengers in microscopic delay management: A game theoretical study”. In: *Transportation Science* 54.3, pp. 785–822. ISSN: 15265447. DOI: 10.1287/trsc.2020.0979. URL: https://pubsonline.informs.org/doi/abs/10.1287/trsc.2020.0979.

Miristice, Lory Michelle Bresciani et al. (2023). “Real-Time passengers forecasting in congested transit networks considering dynamic service disruptions and passenger count data”. In: *2023 8th International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2023*. Institute of Electrical and Electronics Engineers Inc., p. 1. ISBN: 9781665455305. DOI: 10.1109/MT-ITS56129.2023.10241550.

Wang, Xiaohan, Jian Gang Jin, and Lijun Sun (2022). “Real-time dispatching of operating buses during unplanned disruptions to urban rail transit system”. In: *Transportation Research Part C: Emerging Technologies* 139. ISSN: 0968090X. DOI: 10.1016/j.trc.2022.103696.

Yin, Jiateng et al. (2022). “Quantitative analysis for resilience-based urban rail systems: A hybrid knowledge-based and data-driven approach”. In: *Reliability Engineering and System Safety* 219, p. 108183. ISSN: 0951-8320. DOI: 10.1016/J.RESS.2021.108183.

Auld, Joshua et al. (2020). “A stated-preference intercept survey of transit-rider response to service disruptions”. In: *Public Transport* 12.3, pp. 557–585. ISSN: 16137159. DOI: 10.1007/S12469-020-00243-Z/FIGURES/13. URL: https://link.springer.com/article/10.1007/s12469-020-00243-z.

Liu, Yu et al. (2022a). "An intelligent method for accident reconstruction involving car and e-bike coupling automatic simulation and multi-objective optimizations". In: *Accident Analysis & Prevention* 164, p. 106476. ISSN: 0001-4575. DOI: 10.1016/J.AAP.2021.106476.

Rettore, Paulo H. et al. (2019). "Vehicular Data Space: The Data Point of View". In: *IEEE Communications Surveys and Tutorials* 21.3, pp. 2392–2418. ISSN: 1553877X. DOI: 10.1109/COMST.2019.2911906.

SUWANNO, Piyapong et al. (2021). "Impact of Rainfall Intensity on Macroscopic Traffic Variables of Urban Roads Using Data from Bluetooth Detectors". In: *Journal of Human and Environmental Symbiosis* 37.1, pp. 13–23. ISSN: 1346-3489. DOI: 10.32313/JAHES.37.1_13.

Rusmawati, Yanti and Rita Rismala (2016). "Towards the modelling of dynamic train systems: Disruption pattern analysis". In: *2016 4th International Conference on Information and Communication Technology, ICoICT 2016*, pp. 1–5. DOI: 10.1109/ICoICT.2016.7571923.

Zhang, Xiaoling et al. (2016). "An incident database for improving metro safety: The case of shanghai". In: *Safety Science* 84, pp. 88–96. ISSN: 18791042. DOI: 10.1016/j.ssci.2015.11.023.

Louie, Jacob, Amer Shalaby, and Khandker Nurul Habib (2017). "Modelling the impact of causal and non-causal factors on disruption duration for Toronto's subway system: An exploratory investigation using hazard modelling". In: *Accident Analysis & Prevention* 98, pp. 232–240. ISSN: 00014575. DOI: 10.1016/j.aap.2016.10.008.

Grigorev, Artur et al. (2022b). "Traffic incident duration prediction via a deep learning framework for text description encoding". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC* 2022-9, pp. 1770–1777. DOI: 10.1109/ITSC55140.2022.9921768. arXiv: 2209.08735.

Bell, Michael G.H. et al. (2017). "Investigating transport network vulnerability by capacity weighted spectral analysis". In: *Transportation Research Part B: Methodological* 99, pp. 251–266. ISSN: 0191-2615. DOI: 10.1016/J.TRB.2017.03.002.

Jiang, Ruoyun, Qing Chang Lu, and Zhong Ren Peng (2018). "A station-based rail transit network vulnerability measure considering land use dependency". In: *Journal of Transport Geography* 66, pp. 10–18. ISSN: 0966-6923. DOI: 10.1016/J.JTRANGEO.2017.09.009.

Chen, Hongyu et al. (2021). "Simulation-based vulnerability assessment in transit systems with cascade failures". In: *Journal of Cleaner Production* 295, p. 126441. ISSN: 0959-6526. DOI: 10.1016/J.JCLEPRO.2021.126441.

Cao, Na and Haiqing Cao (2020). "Exploring the robustness of urban bus network: A case from Southern China". In: *Chinese Journal of Physics* 65, pp. 389–397. ISSN: 0577-9073. DOI: 10.1016/J.CJPH.2020.03.012.

Wang, Jianhong et al. (2021a). "The strong precipitation of the dry warm front cyclone in syria and its prediction by data mining modeling". In: *Atmosphere* 12.12. ISSN: 20734433. DOI: 10.3390/atmos12121667. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85121995898&doi=10.3390%2Fatmos12121667&partnerID=40&md5=3bfeca38410e101d81bf3e537c92141e.

Cumelles, Joel, Oriol Lordan, and Jose M. Sallan (2021). "Cascading failures in airport networks". In: *Journal of Air Transport Management* 92, p. 102026. ISSN: 0969-6997. DOI: 10.1016/J.JAIRTRAMAN.2021.102026.

Ma, Junfeng et al. (2021). "Scheduling-measure dependent modelling of delay propagation on a single railway line: A matrix transformation approach". In: *IET Intelligent Transport Systems* 15.11, pp. 1429–1439. ISSN: 1751-9578. DOI: 10.1049/ITR2.12112. URL: https://onlinelibrary.wiley.com/doi/full/10.1049/itr2.12112https://onlinelibrary.wiley.com/doi/abs/10.1049/itr2.12112https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/itr2.12112.

Dong, Shangjia et al. (2020). "Probabilistic modeling of cascading failure risk in interdependent channel and road networks in urban flooding". In: *Sustainable Cities and Society* 62, p. 102398. ISSN: 2210-6707. DOI: 10.1016/J.SCS.2020.102398.

Wu, Yipeng et al. (2021). "Propagation model of cascading failure based on discrete dynamical system". In: *Reliability Engineering & System Safety* 209, p. 107424. ISSN: 0951-8320. DOI: 10.1016/J.RESS.2020.107424.

Tang, Jinjun et al. (2020). "Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review". In: *Analytic Methods in Accident Research* 27, p. 100123. ISSN: 2213-6657. DOI: 10.1016/J.AMAR.2020.100123.

Meng, Lingyun and Xuesong Zhou (2011). "Robust single-track train dispatching model under a dynamic and stochastic environment: A scenario-based rolling horizon solution approach". In: *Transportation Research Part B: Methodological* 45.7, pp. 1080–1102. ISSN: 0191-2615. DOI: 10.1016/J.TRB.2011.05.001.

Zhan, Shuguang et al. (2015). "Real-time high-speed train rescheduling in case of a complete blockage". In: *Transportation Research Part B: Methodological* 78, pp. 182–201. ISSN: 01912615. DOI: 10.1016/j.trb.2015.04.001.

Veelenturf, Lucas P. et al. (2016). "A Railway Timetable Rescheduling Approach for Handling Large-Scale Disruptions". In: *Transportation Science* 50.3, pp. 841–862. ISSN: 1526-5447. DOI: 10.1287/TRSC.2015.0618.

Ghaemi, Nadjla, Oded Cats, and Rob M.P. Goverde (2017). "Railway disruption management challenges and possible solution directions". In: *Public Transport* 9.1-2, pp. 343–364. ISSN: 16137159. DOI: 10.1007/s12469-017-0157-z. URL: https://link-springer-com.ezproxy2.library.usyd.edu.au/article/10.1007/s12469-017-0157-z.

Li, Yang et al. (2021). "Discrete-event simulations for metro train operation under emergencies: A multi-agent based model with parallel computing". In: *Physica A: Statistical Mechanics and its Applications* 573, p. 125964. ISSN: 03784371. DOI: 10.1016/j.physa.2021.125964.

Nielsen, LK Lars Kjaer (2011). "Rolling Stock Rescheduling in Passenger Railways Applications in Short-term Planning and in Disruption Management". PhD thesis. Erasmus University Rotterdam. ISBN: 978-90-5892-267-0. URL: www.b-en-t.nl.

Melo, Patricia C. et al. (2011). "Determinants of delay incident occurrence in urban metros". In: *Transportation Research Record* 2216, pp. 10–18. ISSN: 03611981. DOI: 10.3141/2216-02.

Cadarso, Luis, Ángel Marín, and Gábor Maróti (2013). "Recovery of disruptions in rapid transit networks". In: *Transportation Research Part E: Logistics and Transportation Review* 53.1, pp. 15–33. ISSN: 13665545. DOI: 10.1016/j.tre.2013.01.013. URL: https://www.sciencedirect.com/science/article/pii/S1366554513000239.

Schipper, Danny and Lasse Gerrits (2018). "Differences and similarities in European railway disruption management practices". In: *Journal of Rail Transport Planning and Management* 8.1, pp. 42–55. ISSN: 2210-9706. DOI: 10.1016/J.JRTPM.2017.12.003.

Luo, Chunling and Lei Xu (2021). "Railway disruption management: Designing bus bridging services under uncertainty". In: *Computers and Operations Research* 131, p. 105284. ISSN: 0305-0548. DOI: 10.1016/J.COR.2021.105284.

Qi, Liang, Meng Chu Zhou, and Wen Jing Luan (2018). "A Two-level Traffic Light Control Strategy for Preventing Incident-Based Urban Traffic Congestion". In: *IEEE Transactions on Intelligent Transportation Systems* 19.1, pp. 13–24. ISSN: 15249050. DOI: 10.1109/TITS.2016.2625324.

Mihăiţă, Adriana Simona, Laurent Dupont, and Mauricio Camargo (2018). "Multi-objective traffic signal optimization using 3D mesoscopic simulation and evolutionary algorithms". In: *Simulation Modelling Practice and Theory* 86, pp. 120–138. ISSN: 1569-190X. DOI: 10.1016/J.SIMPAT.2018.05.005. URL: https://www.sciencedirect.com/science/article/pii/S1569190X18300686.

Mao, Tuo et al. (2021). "Boosted Genetic Algorithm Using Machine Learning for Traffic Control Optimization". In: *IEEE Transactions on Intelligent Transportation Systems*. DOI: 10.1109/TITS.2021.3066958. URL: https://ieeexplore.ieee.org/abstract/document/9408242.

He, Yacui et al. (2019). "Optimizing Bus Bridging Routing in Response to Disruptions of Urban Transit Rail". In: *Proceedings of the 31st Chinese Control and Decision Conference, CCDC 2019*, pp. 3839–3842. DOI: 10.1109/CCDC.2019.8832746.

Yang, Xin et al. (2019). "An energy-efficient rescheduling approach under delay perturbations for metro systems". In: *Transportmetrica B* 7.1, pp. 386–400. ISSN: 21680582. DOI: 10.1080/21680566.2017.1421109. URL: https://www.tandfonline.com/doi/abs/10.1080/21680566.2017.1421109.

Dobler, Christoph and Kai Nagel (2016). "Within-Day Replanning". In: *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press. Chap. 30, pp. 187–200. DOI: 10.5334/baw.30.

Zhang, Shuyang et al. (2020). "Metro system disruption management and substitute bus service: a systematic review and future directions". In: *Transport Reviews* 41.2, pp. 230–251. ISSN: 01441647. DOI: 10.1080/01441647.2020.1834468. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85092787924&doi=10.1080%2F01441647.2020.1834468&partnerID=40&md5=b93e56430c812bab8d9c64c6bc42a973https://www.tandfonline.com/doi/abs/10.1080/01441647.2020.1834468.

Balcombe, Richard et al. (2004). "The demand for public transport: a practical guide". In: *Transport Policy* 13, pp. 295–306. ISSN: 0967070X. DOI: 10.1016/j.tranpol.2005.12.004.

Wismans, L. J.J. et al. (2018). "Improving A Priori Demand Estimates Transport Models using Mobile Phone Data: A Rotterdam-Region Case". In: *Journal of Urban Technology* 25.2, pp. 63–83. ISSN: 14661853. DOI: 10.1080/10630732.2018.1442075. URL: https://doi.org/10.1080/10630732.2018.1442075.

Ta, Viet Cuong et al. (2018). "Smartphone-Based User Positioning in a Multiple-User Context with Wi-Fi and Bluetooth". In: *IPIN 2018 - 9th International Conference on Indoor Positioning and Indoor Navigation* September, pp. 206–212. DOI: 10.1109/IPIN.2018.8533809. arXiv: 1807.05716.

Huang, Zhiren, Alonso Espinosa Mireles De Villafranca, and Charalampos Sipetas (2022). “Sensing Multi-modal Mobility Patterns: A Case Study of Helsinki using Bluetooth Beacons and a Mobile Application”. In: *Proceedings - 2022 IEEE International Conference on Big Data, Big Data 2022*, pp. 2007–2016. DOI: 10.1109/BigData55660.2022.10020578. arXiv: 2209.13537.

Rubio-Herrero, Javier and Jesús Muñuzuri (2021). “Indirect estimation of interregional freight flows with a real-valued genetic algorithm”. In: *Transportation* 48.1, pp. 257–282. ISSN: 15729435. DOI: 10.1007/s11116-019-10050-6. URL: https://doi.org/10.1007/s11116-019-10050-6.

Ai, Xinbo (2017). “Node Importance Ranking of Complex Networks with Entropy Variation”. In: *Entropy 2017, Vol. 19, Page 303* 19.7, p. 303. ISSN: 1099-4300. DOI: 10.3390/E19070303. URL: https://www.mdpi.com/1099-4300/19/7/303/htmhttps://www.mdpi.com/1099-4300/19/7/303.

Qi, Xiao et al. (2021). “Autonomous Vehicles’ Car-Following Drivability Evaluation Based on Driving Behavior Spectrum Reference Model:” in: *Transportation Research Board* 2675.7, pp. 129–141. ISSN: 21694052. DOI: 10.1177/0361198121994857. URL: https://journals.sagepub.com/doi/10.1177/0361198121994857.

Wei, Jie et al. (2022). “Nonlinear Model-Based Subway Station-Level Peak-Hour Ridership Estimation Approach in the Context of Peak Deviation”. In: *Transportation Research Board* 2676.6, pp. 549–564. DOI: 10.1177/03611981221075624.

Suprayitno, Hitapriya (2018). “Searching the Correct and Appropriate Deterrence Function General Formula for Calculating Gravity Trip Distribution Model”. In: *IPTEK The Journal of Engineering* 4.3. DOI: 10.12962/joe.v4i3.3762.

Shannon, C. E. (1948). “A Mathematical Theory of Communication”. In: *Bell System Technical Journal* 27.3, pp. 379–423. ISSN: 15387305. DOI: 10.1002/J.1538-7305.1948.TB01338.X.

McClean, Sally I. (2003). “Data Mining and Knowledge Discovery”. In: *Encyclopedia of Physical Science and Technology*, pp. 229–246. DOI: 10.1016/B0-12-227410-5/00845-0.

Nie, Tingyuan et al. (2016). “Using mapping entropy to identify node centrality in complex networks”. In: *Physica A: Statistical Mechanics and its Applications* 453, pp. 290–297. ISSN: 0378-4371. DOI: 10.1016/J.PHYSA.2016.02.009.

Australian Bureau of Statistics (2021). *Australian Bureau of Statistics*. URL: https://www.abs.gov.au/ (visited on 08/21/2021).

Xiao, Yao et al. (2016). “A pedestrian flow model considering the impact of local density: Voronoi diagram based heuristics approach”. In: *Transportation Research Part C: Emerging Technologies* 68, pp. 566–580. ISSN: 0968-090X. DOI: 10.1016/J.TRC.2016.05.012.

Nikolić, Marija and Michel Bierlaire (2018). “Data-driven spatio-temporal discretization for pedestrian flow characterization”. In: *Transportation Research Part C: Emerging Technologies* 94, pp. 185–202. ISSN: 0968-090X. DOI: 10.1016/J.TRC.2017.08.026.

OpenData (2017). *Open Data | TfNSW Open Data Hub and Developer Portal*. (Visited on 12/12/2017).

TransitFeeds (2017). *Greater Sydney GTFS - TransitFeeds*. (Visited on 12/12/2017).

Thorsen, Inge and Jens Petter Gitlesen (1998). “Empirical evaluation of alternative model specifications to predict commuting flows”. In: *Journal of Regional Science* 38.2, pp. 273–292. ISSN: 00224146. DOI: 10.1111/1467-9787.00092.

Grange, Louis de, Enrique Fernández, and Joaquín de Cea (2010). "A consolidated model of trip distribution". In: *Transportation Research Part E: Logistics and Transportation Review* 46.1, pp. 61–75. ISSN: 13665545. DOI: 10.1016/j.tre.2009.06.001. URL: http://dx.doi.org/10.1016/j.tre.2009.06.001.

Saberi, Meead et al. (2018). "Understanding the impacts of a public transit disruption on bicycle sharing mobility patterns: A case of Tube strike in London". In: *Journal of Transport Geography* 66, pp. 154–166. ISSN: 0966-6923. DOI: 10.1016/J.JTRANGEO.2017.11.018.

Yuming Ou and Adriana-Simona Mihăiţă and Fang Chen (2022). "Big data processing and analysis on the impact of COVID-19 on public transport delay". In: *Data Science for COVID-19*. Ed. by Utku Kose Khanna et al., pp. 257–278. ISBN: 978-0-323-90769-9. DOI: https://doi.org/10.1016/B978-0-323-90769-9.00010-4. URL: https://www.sciencedirect.com/science/article/pii/B9780323907699000104.

Zhao, Dong (2019). "Bus Bunching Modelling and Control: A Passenger-oriented Approach". Master Dissertation. University of Sydney. URL: https://ses.library.usyd.edu.au/handle/2123/23899.

Assemi, Behrang et al. (2020). *Improving alighting stop inference accuracy in the trip chaining method using neural networks*. 1. Springer, pp. 89–121. ISBN: 1246901900. DOI: 10.1007/s12469-019-00218-9.

Saleh, Khaled, Artur Grigorev, and Adriana Simona Mihaita (2022). "Traffic Accident Risk Forecasting using Contextual Vision Transformers". In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC* 2022-Octob, pp. 2086–2092. DOI: 10.1109/ITSC55140.2022.9921978. arXiv: 2209.11180.

Grigorev, Artur et al. (2022a). "Incident duration prediction using a bi-level machine learning framework with outlier removal and intra–extra joint optimisation". In: *Transportation Research Part C: Emerging Technologies* 141.May, p. 103721. ISSN: 0968090X. DOI: 10.1016/j.trc.2022.103721. arXiv: 2205.05197. URL: https://doi.org/10.1016/j.trc.2022.103721.

Mihaita, Adriana-Simona et al. (2019). "Arterial incident duration prediction using a bi-level framework of extreme gradient-tree boosting". In: *Proc. of ITS World Congress (ITSWC 2019), Singapore* October, pp. 21–25. arXiv: 1905.12254. URL: http://arxiv.org/abs/1905.12254.

Shafiei, S. and Mihaita, A.S. and Cai, C. (2019). "Demand Estimation and Prediction for Short-term Traffic Forecasting in Existence of Non-recurrent Incidents". In: *Proc. of ITS World Congress (ITSWC 2019), Singapore*.

Shafiei, Sajjad; et al. (2020). "Short-term traffic prediction under non-recurrent incident conditions integrating data-driven models and traffic simulation". In: *Transportation Research Board (TRB) Annual Meeting*. URL: http://hdl.handle.net/10453/138721.

Mao, Tuo, Adriana-Simona Mihaita, and Chen Cai (2019). "Traffic signal control optimization under severe incident conditions using Genetic Algorithm". In: October, pp. 21–25. arXiv: 1906.05356. URL: http://arxiv.org/abs/1906.05356.

Abera, Yosef and Dereje Hailemariam (2018). "Spatio-temporal Mobile Data Traffic Modeling Using Fourier Transform Techniques". In: *9th International Conference on Information and Communication Technology Convergence: ICT Convergence Powered by Smart Intelligence, ICTC 2018*, pp. 20–24. DOI: 10.1109/ICTC.2018.8539698.

Chindanur, Narendra Babu and Pallaviram Sure (2018). "Low-dimensional models for traffic data processing using graph fourier transform". In: *Computing in Science and Engineering* 20.2, pp. 24–37. ISSN: 15219615. DOI: 10.1109/MCSE.2018.110111913.

Brunton, Steven L. and J. Nathan Kutz (2021). *Data-driven science and engineering: machine learning, dynamical systems, and control*. Cambridge University Press, p. 614. ISBN: 9781009098489, 1009098489. DOI: 10.1109/mcs.2021.3076544.

Online-supplement (2022). *Appendix: Traffic disruption modelling with mode shift in multi-modal networks*. URL: http://www.simonamihaita.com/papers/2022_IEEE_ITSC_Initial_Multi_modal_APPENDIX_DongZhao.pdf.

Rahman, Md Mokhlesur et al. (2021). "Traffic congestion and its urban scale factors: Empirical evidence from American urban areas". In: *International Journal of Sustainable Transportation*. ISSN: 15568334. DOI: 10.1080/15568318.2021.1885085. URL: https://www.tandfonline.com/doi/full/10.1080/15568318.2021.1885085.

Ettema, Dick et al. (2016). "Travel Mode Use, Travel Mode Shift and Subjective Well-Being: Overview of Theories, Empirical Findings and Policy Implications". In: *Mobility, Sociability and Well-Being of Urban Living*, pp. 129–150. DOI: 10.1007/978-3-662-48184-4_7. URL: https://link.springer.com/chapter/10.1007/978-3-662-48184-4_7.

Ma, Liang, Corinne Mulley, and Wen Liu (2017). "Social marketing and the built environment: What matters for travel behaviour change?" In: *Transportation* 44.5, pp. 1147–1167. ISSN: 15729435. DOI: 10.1007/s11116-016-9698-2. URL: https://link.springer.com/article/10.1007/s11116-016-9698-2.

Adler, Martin W. and Jos N. van Ommeren (2016). "Does public transit reduce car travel externalities? Quasi-natural experiments' evidence from transit strikes". In: *Journal of Urban Economics* 92, pp. 106–119. ISSN: 0094-1190. DOI: 10.1016/J.JUE.2016.01.001.

Nguyen-Phuoc, Duy Q. et al. (2018). "Congestion relief and public transport: An enhanced method using disaggregate mode shift evidence". In: *Case Studies on Transport Policy* 6.4, pp. 518–528. ISSN: 2213-624X. DOI: 10.1016/J.CSTP.2018.06.012. URL: https://www.sciencedirect.com/science/article/pii/S2213624X17303668?via%3Dihub.

Das, Sanhita et al. (2021). "Impact of COVID-19: A radical modal shift from public to private transport mode". In: *Transport Policy* 109, pp. 1–11. ISSN: 0967-070X. DOI: 10.1016/J.TRANPOL.2021.05.005. URL: https://www.sciencedirect.com/science/article/pii/S0967070X21001438.

Guzman, Luis A. et al. (2021). "Revisiting the Benefits of Combining Data of a Different Nature: Strategic Forecasting of New Mode Alternatives". In: *Journal of Advanced Transportation* 2021. ISSN: 20423195. DOI: 10.1155/2021/6672961. URL: https://www.hindawi.com/journals/jat/2021/6672961/.

Aftabuzzaman, Md, Graham Currie, and Majid Sarvi (2010). "Modeling the Spatial Impacts of Public Transport on Traffic Congestion Relief in Melbourne, Australia:" in: *https://doi.org/10.3141/2144-01* 2144, pp. 1–10. ISSN: 03611981. DOI: 10.3141/2144-01. URL: https://journals.sagepub.com/doi/10.3141/2144-01.