Università di Genova

DOCTORAL THESIS

---

# Dynamic traffic assignment models for disrupted networks

---

*Author:*
Enrico SIRI

*Supervisor:*
Dr. Silvia SIRI

*A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy*

*in the*

Transport Systems Engineering Lab
Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS)

*"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."*

Dave Barry

UNIVERSITY OF GENOVA

# *Abstract*

Department of Informatics, Bioengineering, Robotics and Systems Engineering (DIBRIS)

Doctor of Philosophy

**Dynamic traffic assignment models for disrupted networks**

by Enrico SIRI

Transportation infrastructure systems are one of the cornerstones on which modern societies are founded. They allow the movement of people and goods by enabling business activities, the setting up of supply chains, and they provide access to vital resources and services. It is commonly believed that due to their vast scale and complexity, transportation systems are among the most vulnerable infrastructures in the occurrence of a disruption, i.e. an event that involves extensive damage to people or physical facilities.

The growing awareness about this issue in recent years has led to a growing body of literature on the topic of performance evaluation of transportation networks when affected by disruptive events, aimed at providing adequate estimations of network operability in such contexts. A peculiarity of transportation is to be a socio-technical system where the transportation supply, represented by the infrastructure and related services, interacts with the transportation demand, consisting of all those individuals who access the infrastructure at any given time. This property makes such systems inherently complex and consequently an analysis of their vulnerability in the face of disruptive events needs to be able to account for these interactions.

The aim of the present thesis is therefore to develop methodologies capable of convincingly portraying the reaction of such systems in the face of disruptive events by modeling the dynamics that emerges between users and the infrastructure. It is reasonable to assume that travelers due to changed system conditions will adapt their behavior to some degree in order to mitigate the consequences of such events. In this regard, three modeling approaches are presented in this manuscript to address the need to represent this reaction phenomenon.

The first approach involves the use of an inter-period traffic assignment model able to represent the evolution of users' mobility choices in a dynamic context. For each period, the users' reaction is estimated by solving an assignment model thus computing the optimal flow distribution given the current congestion conditions. User habits are taken into account by appropriately limiting the extent of flow redistributions in order to represent the gradual adaptation of the system to the new situation.

Large perturbations can trigger modal shift phenomena between one transport sub-system and another. In this regard, a multi-modal multi-class scenario analysis model is then presented. Railway and road transport sub-networks are thus embedded into an extended hyper-network to model flow exchanges between this two sub-systems. Class-specific assignment models are employed to determine the choice behavior for passenger flows and freight flows. The results of these choices are then routed through the network by means of a discrete-time dynamic flow model.

Finally, the idea that users' behavior may be influenced by their habits is further explored within a path-based inter-period assignment model. It is suggested that users' route choice process is not only influenced by the travel costs of available alternatives but also by users' familiarity with them. More specifically, if changing traffic conditions suddenly make a specific route disadvantageous, users will tend to prefer those that are most topologically similar to the one they are abandoning. This assumption is then investigated by demonstrating that it implies considering a rationally bounded user choice process. The steady state reached by the system as a result of the equilibration process is then detailed and a rigorous proof is provided to show that it is equivalent to a Boundedly Rational User Equilibrium.

All three approaches has been successfully applied on appropriate test networks where a disruption is simulated by altering the network topology.

These models can provide an important contribution to transportation network vulnerability and resilience analyses willing to take into account the interaction between the infrastructure and users.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| **AIMSUN** | **A**dvanced **I**nteractive **M**icroscopic **S**imulator for **U**rban and Non Urban **N**etworks |
| **BR** | **B**ounded **R**ationality |
| **BRUE** | **B**ounded **R**ationality **U**ser **E**quilibrium |
| **CONTRAM** | **Con**tinous **Tr**affic **A**ssignment **M**odel |
| **CTM** | **C**ell **T**ransmission **M**odel |
| **DRAP** | **D**iscrete **R**ational **A**djustment **P**rocess |
| **DTA** | Fully-**D**ynamic **T**raffic **A**ssignment |
| **DTD** | **D**ay-**t**o-**D**ay Traffic Assignment |
| **DUE** | **D**ynamic **U**ser **E**quilibrium |
| **DynaMIT** | **D**ynamic **N**etwork **A**ssignment for the **M**anagement of **I**nformation to **T**ravelers |
| **DYNASMART** | **Dy**namic **N**etwork **A**ssignment-**S**imulation **M**odel for **A**dvanced **R**oad **T**elematics |
| **EEA** | **E**uropean **E**nvironment **A**gency |
| **GLTM** | **G**eneral **L**ink **T**ransmission **M**odel |
| **HCM** | **H**ighway **C**apacity **M**anual |
| **IUE** | **I**nertial **U**ser **E**quilibrium |
| **LTM** | **L**ink **T**ransmission **M**odel |
| **LWR** | **L**ightill-**W**hitham-**R**ichards traffic flow model |
| **METANET** | Highway Network Traffic Flow Model (**M**odéle d′ **E**coulement du **T**raffic **A**utoroutièr **N**etwork) |
| **MFD** | **M**acroscopic **F**oundamental **D**iagram |
| **MSA** | **M**ethod of **S**uccessive **A**verages |
| **OD** | **O**rigin **D**estination pair |
| **PR** | **P**erfect **R**ationality |
| **PW** | **P**ayne-**W**ithman traffic flow model |

| | |
|---|---|
| **SATURN** | **S**imulation and **A**ssignment of **T**raffic to **U**rban **R**oad **N**etworks |
| **SO** | **S**ystem **O**ptimum |
| **STA** | **S**tatic **T**raffic **A**ssignment |
| **SUE** | **S**tochastic **U**ser **E**quilibrium |
| **SUMO** | **S**imulation of **U**rban **Mo**bility |
| **TTC** | **T**otal **T**ravel generalized **C**ost |
| **TTT** | **T**otal **T**ravel **T**ime |
| **UE** | Deterministic **U**ser **E**quilibrium |
| **VISSIM** | Traffic in cities - simulation model (**V**erkehr **I**n **S**tädten - **Sim**ulationsmodell) |

# Chapter 1

# Introduction

Our societies and economies rely on a number of critical infrastructure systems in the absence of which they simply could not operate. These systems ensure the supply of a variety of energy sources (e.g. power grids, gas pipelines), the continuous provision of essential resources (e.g. water supply), the information transferability and storage (e.g. information and communication systems) and the mobility of people and goods (e.g. transport systems), among others. All these systems have grown in scope and complexity over time to serve the needs of an increasingly dynamic and global society. Countless other higher level systems rely on these physical infrastructures such as supply chains, finance and healthcare, social and knowledge networks to name a few. The increase in complexity has made this fundamental systems more vulnerable and their highly interdependent nature implies that a failure of any of them has the potential to reverberate across the others.

Unfortunately, out of similar and equally crucial systems transport networks are among the most suffered infrastructure systems in exogenous (e.g. natural disasters) or endogenous (e.g. accidents, technical failures) disruptions. Network failure can have more or less severe consequences depending on the nature of the disruption. Interrupted road sections, trains break down or flight cancellations are all events that imply an immediate increase in travel time for people and goods and have a direct or indirect impact in terms of economic and social costs, unfortunately also extremely severe. Since the transportation network is one of the backbones of our society, it is not surprising that considerable research has developed over the years aiming to understand the mechanisms of its systemic functioning and the underlying reasons for its vulnerability in order to identify, predict and possibly alleviate this issue. This challenge is greatly hardened because transportation systems are inherently complex due to the fact that they consist of several different elements interacting in a variety of ways where only a fraction can be considered "technical" and thus governed by physical laws peculiar to more classical engineering disciplines. As a consequence, the local manifestation of a disruption may alter the normal operation level of a transportation network far beyond its epicenter and the propagation dynamics of the effects may be complex and counterintuitive.

Several frameworks and concepts have been developed to investigate the performance evolution of a transportation system when exposed to various types of disturbances from day-to-day fluctuations to actual large-scale disasters. Among these are reliability analysis, robustness and vulnerability analysis and resilience analysis. The effectiveness of these methodologies, taking into account their peculiarities, can vary considerably due to the amount of information available for the analysis. A conventional approach, as will be illustrated in this chapter, is to rely on a representation, more or less abstract depending on the context, of the transportation network and, based on its topology, the analysis is performed and the corresponding metrics are estimated. But as mentioned, on its own, the topology of a transport network represents a partial information. Equally important is the behavior that emerges as a result of the combined choices of a multitude of individuals who at any given time have access to the transportation infrastructure.

In light of these considerations, the purpose of this thesis is to provide models able to convincingly represent the complex interaction that takes place between users and the transport system. A transportation infrastructure is by definition a "scarce" resource, in the sense that at any given time it is able to provide a limited level of performance. The more people who want to exploit the resource at a given time, the more the cost of the resource increases. The cost to be paid in this case is a travel cost, often travel time, and the scarcity of the transportation resource takes the form of "congestion" or " disservice". Given a transport infrastructure and a mobility demand, an implicit trade-off between all users is then reached. In the transportation science, this trade-off is called *equilibrium*. The main objective of this thesis is therefore to represent the equilibration process that arises when the transportation infrastructure undergoes a significant variation in the level of service caused by a disruption. Such a perturbations alters the system status quo and it is reasonable to assume that users will react accordingly by adapting their choices in the immediate future, for example by changing their usual home-to-work route or choosing a different transportation mode. Representing this phenomenon can provide valuable information for developing a wide variety of network analyses.

The rest of the chapter is organized as follows. In section 1.1 the issue of network disruption analyses is introduced with particular attention for transportation networks. The various approaches and the related available metrics for assessing the state of the network are presented. Following, section 1.2 is devoted to introducing the existing approaches for estimating the necessary data required to formulate the analyses. In particular, the use of traffic assignment techniques is discussed in detail. Finally, in section 1.3, the overall structure of the present thesis is summarized.

## 1.1 Network-disruption analysis

Network-disruption analysis embodies a multitude of methodological approaches that have been widely used in a variety of contexts including transportation planning and infrastructure maintenance when operating at a subnormal functional level. The main objective is to identify the most critical portions of a network, i.e. those links or nodes that induce serious repercussions on the entire network (Sullivan, Aultman-Hall, and Novak, 2009). Network disruption analysis is a research area that has seen a significant increase in attention in recent years, sadly due to a series of disasters that have impacted businesses, infrastructure and entire communities such as the earthquake of Kobe in Japan (1995), the attack on the World Trade Center (2001), the I-35 Minneapolis bridge collapse (2007), the earthquake in Tōhoku Japan (2011) and the Polcevera Viaduct collapse in Genoa (2018), to name a few. In the Technical Report 13/2010 of the European Environment Agency (EEA) titled "Mapping the impacts of natural hazards and technological accidents in Europe", it is reported that the number and the magnitude of disasters increased in Europe during the decade 1999-2009.

Primarily, it is important to distinguish between the various typologies of disruption because, especially in the case of a transportation network, the assumptions related to the disruptive event are fundamental to the formulation of the analysis results (Sullivan et al., 2010). A preliminary classification can be drawn between internal (endogenous to the system) or external (exogenous to the system) causes and between accidental or

| Cause | Accidental | Intentional |
|---|---|---|
| Internal | Technical failures, mishaps | Labour market conflicts |
| Esternal | Natural disasters | Sabotage, attacks |

TABLE 1.1: Characterization of causes of threats and disruptions to the transport system (Adapted from Mattsson and Jenelius (2015))

intentional disruptions (Mattsson and Jenelius, 2015), as summarized in Table 1.1.

Internally originated disruptions can be induced by human errors (by staff or users), infrastructure over-utilization, malfunctions or complete subsystem failures. They may also be intentional, like sabotage actions instigated as an act of protest resulting from internal labor market conflict for example.

Externally originated disruptions, when unintentional, can be a consequence of natural disasters of various kinds such as earthquakes, tsunamis, hurricanes, wildfires, storm rains, to name a few. Conversely, when intentional, an external threat may be embodied in an actual infrastructure attack such as sabotage, a terrorist attack or an act of war. As a matter of fact, transportation systems, considering how crucial they are in the daily lives of people and their activities, are often the perfect target of all the actions whose goal is to provoke as much damage as possible to a community. This class of disruptions is harder to handle than the others since it is extremely complex to build valid predictive models able to determine the place and time of an intentional attack while in other contexts, such as natural disasters, it is still possible despite several difficulties to collect statistically relevant data to be used for predictive and hopefully prevention purposes.

Hasan and Foliente (2015) propose to classify disruptive events according to two dimensions: the time available to prepare for the event (related to its predictability) and the duration of the event (related to its magnitude). Therefore, the depth and sophistication of the analysis will depend on the amount of time available, while model features and period of study will depend on the type of disruption and the estimated duration. The duration of a disruption in turns depends not only on the span of its physical manifestation (e.g. few seconds for an earthquake or few days for a flood) but also on the duration of the post-event recovery phase resulting from the interaction between the event itself and the local systems and communities involved. Furthermore, the authors emphasize the centrality of both the magnitude and the probability associated with the event. Many disruptive events are associated with low occurrence probability but at the same time with potentially devastating effects. Within the context of risk analysis, the authors point out that in recent years the researchers' effort has largely focused on the study of those events that are extremely rare ("black swan") but associated with severe consequences.

Finally in the context of transport networks, Calvert and Snelder, 2018 emphasize the difference between local and global disturbances where the former have a limited range while the latter have the potential to involve extensive portions of the network. However, distinguishing in advance between the two is not trivial. Events all in all limited may result in extensive cascading effects because of the complex interdependence between infrastructure and transportation demand.

The performance of a transportation network in relation to the occurrence of a disruptive event has been studied employing multiple different methodologies and as many different concepts. They can be clustered into three macro groups according to the time scale, the magnitude and the duration of network performance fluctuations: (1) the concept of *reliability* is associated with frequent but contained fluctuations; (2)the concept of *vulnerability* and its complementary *robustness* are associated to significant but rare performance fluctuations lies ; (3) the newer concept of *resilience* is related to both the magnitude and the duration of the fluctuations lies.

### 1.1.1 Reliability

The concept of reliability is commonly employed in risk analysis and refers to "the probability that a system will operate adequately for the time period intended under the operating conditions encountered" (Wakabayashi and Iida, 1992). In the transportation domain, reliability is associated with the concepts of stability, certainty and predictability of travel conditions and thus intrinsically linked to probability theory. As Berdica (2002) pointed out, reliability analysis almost exclusively consist of likelihood estimation. That is because the phenomena under study are associated with daily and relatively minute variations in travel conditions. In other words, the focus is on those events that occur extremely frequently but do not involve major system performance degradations but rather frequent and localized fluctuations. A transportation system, or a portion of it, is therefore defined as *unreliable* if it is prone to frequent performance fluctuations. Methodologically, reliability is related to the standard deviation, variance, or some other dispersion measures estimated for the statistical distribution of travel times or generalized travel costs according to a specified perturbation (Bell, 1999; Tu, Lint, and Zuylen, 2008; Jong and Bliemer, 2015; Taylor, 2013; Zheng et al., 2018).

Although it is not always easy to draw a clear line, reliability is a fundamentally different concept compared to vulnerability and resilience, related mostly to rare events. In reliability analysis, it is relatively easy to collect data on travel time or cost variations due to the high frequency of the events. Consequently, establishing appropriate probability distributions related to such events is quite reasonable. Similarly, a traveler who is subject to these kinds of fluctuations is also able to gain sufficient experience and adapt his or her behavior accordingly. As suggested by Nicholson et al. (2003), it is therefore legitimate to assume that traveler's choices will be influenced by the reliability of the alternatives available to him. For this reason, multiple methodologies have been developed to represent the travelers' choice behavior when dealing with an unreliable network (Nie, Zhang, and Lee, 2004; Shao, Lam, and Tam, 2006; Jiang, Mahmassani, and Zhang, 2011; Taylor, 2013; Kato et al., 2021).

### 1.1.2 Vulnerability and Robustness

Unlike reliability, several definitions of vulnerability can be found in the literature. Berdica (2002) suggests that "vulnerability in the road transportation system is a susceptibility to incidents that can result in considerable reductions in road network serviceability". Such a definition highlights the fundamental aspects underlying the concept of vulnerability, namely that there is an initial failure or accident that undermines the usual functioning of a system and that the consequences of this may be significant. Extending the vulnerability concept to the entire transport system, Jenelius and Mattsson (2015) define "transport system vulnerability as ... society's risk of transport system disruptions and degradations".

Husdal et al. (2004) distinguish between structural-based vulnerability, natural-based vulnerability and traffic-based vulnerability. Structural-based vulnerability concerns the way the network is built and thus to its functional characteristics not limited to its topology. In other words, everything related to the physical nature of the infrastructure such as the state of the road body, its geometry, slope, presence of tunnels, bottlenecks or bridges is considered within the model formulation. Natural-based vulnerability, on the other hand, deals with the features of the land crossed by the road network. The orographic, seismic, hydrological or climatic properties of the territory are linked to the likelihood of events such as rock falls, earthquakes, floods or heavy rains and snowfall. Lastly, traffic-related vulnerability concerns the relationship between the infrastructure and the mobility demand. It assesses, for example, the ability of the network to handle unexpected spikes of demand and may relate to the system's ability to take quick action after an incident.

Vulnerability analysis deals with potentially catastrophic but also extremely rare phenomena and it should therefore proceed in three directions answering three key questions (Kaplan and Garrick, 1981) "What will happen? What are the chances? What will be the consequences?" whose answers provide a description of the scenario, the associated probability and the consequences. As pointed out by Jenelius, Petersen, and Mattsson (2006), estimating the probability associated with rare events can be extremely problematic due to the scarcity of empirical data. Going further, Taylor and D'este (2003) propose to focus the anlysis exclusively on event repercussions while ignoring the odds .

Complementary to vulnerability lies the more recent concept of robustness, defined as "the ability of the system to cope changes without altering its configuration" (Wieland and Wallenburg, 2012). In the context of road networks, this definition is declined by Snelder, Van Zuylen, and Immers

(2012) as "how well a network, under predefined circumstances, is able to maintain the operation level for which it was originally designed". The concepts of vulnerability and robustness are basically interchangeable at least until the likelihood of the event plays a predominant role in the analysis. The concept of robustness is mostly related to the system's response to disruption and thus the magnitude of its effects. A robust network is therefore able to resist to a disruption effectively while avoiding a significant overall drop in performance.

The research related to vulnerability analysis has been developed over the years on two parallel and barely interacting paths. One is topological vulnerability analysis, which focuses on defining metrics and methodologies derived from graph theory in order to correlate the quality of a network to its topological features omitting its functional properties and the interaction with users The second stream, on the other hand, is system-based vulnerability analysis, which while still relying on a graph representation details its properties by employing the functional characteristics of the real network such as the actual link lengths and estimated travel costs.

**Topological Vulnerability Analysis**

In topological vulnerability analyses, the transport network is represented by an abstract graph consisting of nodes and links. Depending on the granularity level, the graph can be direct or indirect while the links weighted or unweighted. The analysis typically assesses the degree of connectivity and efficiency of the network as well as the relative importance of individual components according to topological measures such as the beetwennes centrality. Network efficiency is defined as the average distance between each nodes while link's beetweeness centrality is represented by the fraction of shortest paths that make use of that link. Latora and Marchiori (2005) use the betweenness centrality to investigate the vulnerability level of a public transportation network against intentional attacks. Applying the methodology to the Boston subway network the authors estimate that damaging the most critical elements results in a loss of network efficiency of nearly 30%. Similarly, Demšar, Špatenková, and Virrantaus (2008) apply the same metric to identify the most critical cuts in the Helsinki road network. Berche et al. (2009) estimate the vulnerability of public transportation networks in 14 major cities when impacted by accidental or intentional disruption. Efficiency drop is then evaluated when network elements are removed randomly (accidental disruption) or according to their betweenness centrality (intentional disruption). The authors show that a strategy based on intentional damages based on betweenness centrality is able to result in significant efficiency losses even when the number of elements involved is relatively small. Duan and Lu (2014) come to a similar result in the context of road networks by pointing out, however, that vulnerability analysis may lead to different assessment even for the same network depending on the granularity of the representation. More recently, Bell et al. (2017) employed spectral analysis to assess the most critical network cuts. A cut link is a link that if removed makes the network disconnected. Besides being computationally efficient, this type of analysis allows the identification of potential bottlenecks for traffic flows. Finally, Li, Rong, and Yan (2019) extends the concept of efficiency to entire areas. An area is represented by a subnetwork and is all the more critical the higher the population density, economic development and criticality of the network elements within it. The criticality of a network element is once again evaluated according to the network efficiency loss in the absence of that element.

**System-based vulnerability analysis**

System-based analysis, while still making use of graph theory and similar metrics, as in the case of topological-based analysis, incorporates into the models other types of data such as the transportation demand and the functional characteristics of the network. The graph itself is generally closer to the real network morphology with nodes and links corresponding to real network components and where link weights are obtained according to roads length or estimated travel times or travel costs on the corresponding road sections. A conventional method of carrying out this analysis is to compare some generalized cost metrics before and after disruption. These metrics typically are based on estimating link travel times, traffic flows, link capacities, availability of alternative routes or the increase of user travel time. Jenelius, Petersen, and Mattsson (2006) suggest the use of the increase in generalized travel costs in conjunction with the estimated portion of unsatisfied demand following a disruption as an indicator of network vulnerability. The network links are then ranked according to their importance while the areas according to their level of exposure, i.e. how serious the consequences may be for the specific area. This approach is then expanded by also taking into consideration the number of alternative roads and the duration of the disruption (Jenelius, 2010). A link is therefore particularly critical when the disruption may be long lasting and there is a lack of alternatives. Sullivan et al. (2010) propose an analysis that also accounts for partial link disruption, i.e. link disruption levels below 100% (equivalent to a link removal) are allowed in order to account for partial capacity losses.

A different approach is proposed by Rupi et al. (2015) where the local and global importance of a link is emphasized. The local importance of a link is deduced from the amount of daily traffic flow crossing it while the global importance, in line with other approaches, is related to the increase in travel time and the amount of unsatisfied demand as a result of disruption.

Taylor, Sekhar, and D'Este (2006) rather propose to assess the vulnerability of a network by exploiting an accessibility index, meaning the ability to reach a given area of the network from the others while also taking into account the time required. Consequently, a node is vulnerable if the partial or total degradation of a small number of other network components degrades its accessibility. Conversely, a link is defined as critical if as a consequence of a substantial performance degradation of it, several nodes experience a drop in accessibility level. Similarly, Chen et al. (2007) develop a model comprising demand estimation, mode choice and route choice behavior in order to compute the equilibrium state reached by the system a sufficient amount of time after the disruption, i.e. when traffic flows distribution is assumed to have return to a sufficiently steady state. Therefore, corresponding travel times are utilized to assess network vulnerability based on an accessibility index.

Somewhat complementary to the concept of accessibility, Taylor et al. (2012) propose the concept of remoteness. The corresponding metric is related to the level of network disjunction, defined in terms of the travel time required to reach each portion of the network from the others. An area is vulnerable if following a disruption its remoteness increases substantially. More recently, García-Palomares et al. (2018), as part of a vulnerability analysis of the Spanish road network, introduced a vulnerability metric based on three different measures of accessibility namely population-weighted travel time, potential accessibility capacity and daily accessibility. The potential accessibility capacity of a destination is defined as the sum of the ratios between the amount of population at the destination and the distance ( in terms of travel time) required to reach it from another node while the daily accessibility of a node is related to the ability to reach other nodes within a certain time range. The significance of a destination is represented by the size of its population and consequently a node with high daily accessibility provides quick access to other densely populated nodes. Finally Almotahari and Yazici (2020) conducting a comparative analysis of several metrics employed for link criticality ranking showed that usually link criticality is strongly correlated with traffic flow under normal operating conditions making this information a valid proxy in various context.

### 1.1.3 Resilience

Initially introduced in the study of ecological systems (Holling, 1973), the concept of resilience has been applied to multiple areas of study from economics to social sciences and in transportation and civil infrastructure among others. Although a multitude of domain-specific definitions exists, the shared idea behind most of them is that resilience should be related to the system's ability to return to an acceptable level of operation after a disruption altered its state (Hosseini, Barker, and Ramirez-Marquez, 2016). In the context of transportation systems, Murray-Tuite (2006) has been one of the first to define resilience and to propose a series of measures including ten attributes, from the ability of the system to maintain a certain level of operation when hit, to the speed at which it is able to return to an acceptable level of performance.

The majority of resilience mode-specific definitions (Liao, Hu, and Ko, 2018; Beiler et al., 2013; Baroud, Barker, Ramirez-Marquez, et al., 2014; Bhavathrathan and Patil, 2015) are mostly based on the same overall concept and the associated analyses quantify resilience from two different perspectives (Zhou, Wang, and Yang, 2019): (1) the ability to maintain an acceptable level of operation when hit (associated with the system's vulnerability and robustness); (2) the amount of time and resources required to retrieve lost performance. As shown in Fig. 1.1, performance loss as a result of disruption consists of two main phases. The earliest one coincides with the disruption duration where generally there is a drastic drop in performance (from $P(t_0)$ to $P(t_1)$). The concepts of system robustness and redundancy are tied to this phase, where the former concerns the system's ability to withstand the event-induced damages while the latter reflects the availability of alternatives. Following, there is a recovery phase when the system regains partial or complete operability (from $P(t_1)$ to $P(t_2)$). Tied to this second phase it is possible to define the properties of resourcefulness and rapidity where the former is related to the availability of resources at the affected sites while the latter assesses the system's ability to exploit those resources

The significant difference between resilience and vulnerability is that, for the former, the time dimension of the phenomenon is also taken into account, and in particular the time taken by the system to at least partially recover its performance. Thus a system may not be robust (hence vulnerable) and suffer a significant performance loss but still be considered resilient if it is able to recover in a sufficiently limited time. Resilience then is a broader system property that incorporates vulnerability within it but is not limited to it.

Similarly to vulnerability analysis, resilience metrics can be divided into different categories: topological metrics, attribute-based metrics, and performance-based metrics.

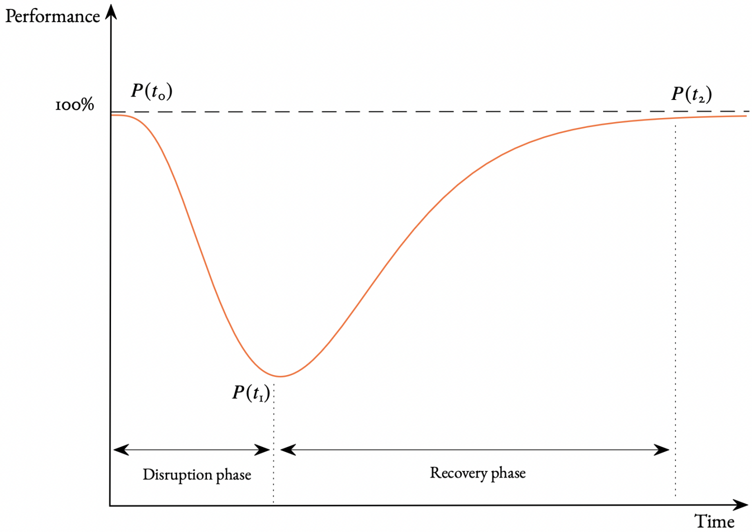



FIGURE 1.1: Two phases of resilience measurement (Adapted from Twumasi-Boakye and Sobanjo (2018))

### Topological Metrics

Topological metrics aim to estimate the resilience of a network by exploiting its structure while ignoring its dynamic properties. Similar to topological vulnerability analysis, the network is represented by an abstract graph whose relative structural properties are then studied. These graph-based properties may be betweenes centrality, efficiency, size of giant component or average shortest path length, where the last two are usually the most commonly employed (Zhou, Wang, and Yang, 2019).

Schintler et al. (2007), implying a raster-based geographic information techniques, estimate network resilience by simulating a series of failures by sequentially removing critical nodes and evaluating the variation in connectivity and shortest path length. Testa, Furtado, and Alipour (2015) apply topological analysis to coastal transportation networks to assess their resilience against particularly adverse climatic conditions. In addition to employing graph-based metrics such as connectivity and average shortest path length, they propose to consider network redundancy, defined in terms of the average number of paths connecting a node to the neighbors of its neighbor. In the context of public transport, Chopra et al. (2016)

assess the resilience of the London Underground network by looking at whether it exhibits small-world properties, namely a network where nodes while poorly directly connected can be indirectly reached in a small number of steps. More recently, Aydin et al. (2018) integrate graph theory with a stress test methodology in order to asses network resilience. Sequentially, the scenario to be represented is determined and the difference in the values assumed by several metrics before and after disruption are computed. Among the employed metrics are efficiency, betwennes centrality, and size of giant components, i.e. the number of nodes within the largest connected subnetwork.

### Attribute-based Metrics

As pointed out previously, the concept of resilience is comprehensive of different partial attributes such as robustness, redundancy, resourcefulness and rapidity (Twumasi-Boakye and Sobanjo, 2018). Attribute-based measures generally focus on one or a subset of these properties, and resilience is estimated based on attribute variations between specific periods. Although robustness and redundancy are certainly quite important features of a network, what distinguishes the concept of resilience from vulnerability is the focus on the system's ability to recover part of the lost performance. For this reason, in this kind of analysis emphasis is placed on the temporal dimension of the phenomenon. Following this approach, Adams, Bekkem, and Toledo-Durán (2012) define appropriate performance reduction and recovery metrics. Referring to Fig. 1.1, the reduction metric is defined as follows

$$\text{Reduction} = \frac{P(t_0) - P(t_1)}{t_1 - t_0} \tag{1.1}$$

where $P(t_0)$ and $P(t_1)$ are the performance before the advent of the disruption and during the initial shock, respectively. Equation (1.1) evaluates the ratio between the net performance loss and the elapsed time. A large performance loss occurring in a short time implies a significant reduction index. Similarly, a recovery metric is also defined as follows

$$\text{Recovery} = \frac{P(t_2) - P(t_1)}{t_2 - t_1} \tag{1.2}$$

where $P(t_2)$ is the level of performance achieved by the system a sufficient amount of time after the advent of disruption. Large performance gains occurring in a short period of time imply a high recovery index. D'Lima and Medda (2015) suggests using the time required for returning to equilibrium after a shock as a metric for assessing the resilience of a system. The authors then apply the methodology to the analysis of the London Underground network portrayed by means of a mean-reversion model, i.e. a system that tends to return to its average operational level after a sufficient amount of time.

Expanding the concept of resilience, Murray-Tuite (2006) recommend employing other metrics to assess additional relevant aspects of a transportation system. Therefore, in addition to a recovery and efficiency metric they also propose a safety metric based on the estimated average number of incidents that occur within the network in a given time frame.

### Performance-based Metrics

Performance-based metrics evaluate the variation in system performance across the whole period during which the effects of the disruption still persist. As a function of the magnitude and duration of performance losses, the resilience is then estimated.

The index of degradation of system quality over time first, firstly proposed by Bruneau et al. (2003), is one of the most widely adopted performance-based resilience metric due to its versatility. Let $Q(t)$ be the index for system's quality at time $t$, $t_0$ the time at which disruption occurs and $\tau$ the duration of effects, resilience index $R$ is calculated based on the total quality loss throughout the whole time horizon.

$$R = \int_{t_0}^{t_0+\tau} \left[100 - Q(t)\right] dt \tag{1.3}$$

The metric defined in (1.3) has therefore been adopted to estimate the resilience of road traffic networks (Bocchini and Frangopol, 2012a; Bocchini and Frangopol, 2012b), rail networks (Adjetey-Bahun et al., 2016), and subways and public transportation networks (Zhu et al., 2016).

Adjetey-Bahun et al. (2016) proposed a modified version of (1.3) where the quality loss is normalized over the length of the disruption period as shown in (1.4).

$$R = \int_{t_0}^{t_0+\tau} \frac{[100 - Q(t)]}{\tau} dt \tag{1.4}$$

The characterization of the quality index depends on the domain in which the analysis is performed. For example within the maritime context, Omer et al. (2012) employ the ratio between travel time before disruption and travel time after disruption between two nodes as an estimator of quality loss as shown in tot.

$$R = \int_{t_0}^{t_0+\tau} \frac{t_{ij(\text{before shock})}}{t_{ij(\text{after shock})}} dt \tag{1.5}$$

A similar approach is proposed by Faturechi and Miller-Hooks (2014) and Bhavathrathan and Patil (2015) in the context of road networks where total travel time is employed as system's quality index.

## 1.2 Data requirements and modeling approaches

The approaches to estimate the response of a network system when affected by a disruptive event exploiting exclusively topological features have the great advantage of requiring limited input data. The resilience or vulnerability assessment is then obtained by designing appropriate scenarios where a series of network components are removed randomly or according to some attack pattern based on some criticality measure. This advantage makes topological metrics-based approaches particularly suitable for the study of large transport network especially when the aim is to inspect the criticality of each network component which necessitates the sequential removal of all links or nodes. However, the simplicity and level of abstraction involved in these approaches also bring significant drawbacks. In reality, it is likely that the consequences of a disruption will depend on several other factors besides the infrastructure topology such as the duration of the event, how many users will be directly affected, how many alternatives they have in terms of destinations, routes or modes but most importantly how they will react. Furthermore, topological approaches still fail to capture the dynamics of the system's response. Beyond its morphology, equally important are the congestion propagation towards network areas not directly involved in the disruption and the behavior of users in the face of these changes.

Approaches employing some kind of performance metric (i.e. system based, attributes based and performance based analysis) represent an attempt to overcome these limitations at the price of requiring much more data. Information associated with transport demand estimation, functional characteristics of the network as well as the formulation of assumptions about user behavior are necessary ingredients for the proper calibration of this type of model.

The approaches for estimating resilience metrics or for vulnerability analysis can be divided into: optimization models, simulation models and data-driven models.

**Optimization models** Optimization models are employed in order to estimate traffic conditions (flow distribution and travel times on the network) through solving appropriate traffic assignment problems (Patriksson, 2015). They are also used for the definition of optimal strategies related to mitigation, preparedness, and recovery operations in response to a disruption (Nair, Avetisyan, and Miller-Hooks, 2010; Faturechi and Miller-Hooks, 2014; Azad, Hassini, and Verma, 2016).

**Simulation models** Simulation models make use of highly accurate representation even down to individual elements (such as vehicles in a road network). They are generally used for reliability evaluations and more generally for modeling the impact of frequent and localized phenomena (Gauthier, Furno, and El Faouzi, 2018). Nevertheless they are sometimes employed for estimating equilibrium traffic conditions (Murray-Tuite, 2006; Osei-Asamoah and Lownes, 2014).

**Data-driven models** Data-driven models, as the name suggests, represent the most data-hungry methodology with the advantage of not having to represent the internal system mechanisms. This class of approaches makes use of a huge amount of appropriately processed empirical data in order to estimate performance variations under different scenarios. Generally, some proxies are used to estimate system properties such as passenger count (D'Lima and Medda, 2015), vehicle counts and speed measure (Adams, Bekkem, and Toledo-Durán, 2012), reconstruction time and cost (Mojtahedi, Newton, and Von Meding, 2017), and shifts in modal choices (Stamos et al., 2015).

### 1.2.1 Traffic Assignment Models

One way to estimate the conditions of a transportation network under a certain scenario is to make use of appropriate traffic assignment models. Belonging to the class of optimization models, a traffic assignment can be defined in the simplest terms as a methodology that allows, given the mobility demand and the functional characteristics of the network, to estimate how users will spread across it on their way to their respective destinations according to some behavioral assumptions. Mobility demand is usually expressed in the form of an origin-destination matrix whose elements represent the amount of flow that seeks to travel between a pair of network nodes. On the other hand, network functional characteristics refer to all the features that make it possible to model the interaction between infrastructure and users such as its topology, capacity and arc length, to name a few (see section 2.2). Essential within the network state estimation process is the set of assumptions related to the users' behavioral model, different assumptions about the underlying motivations involved in the users' choice process lead to the different flow pattern estimation. A standard way of proceeding is to assume that users perform a choice in order to minimize their travel cost or travel time to commute from their origin to their destination (see section 2.3). If users also dispose of accurate information about the network condition, it is reasonable to assume that they will choose the minimum-cost path. Furthermore, if it is assumed that the network has infinite capacity, i.e. it can handle all traffic demand without any degradation in performance level, then a traffic assignment is equivalent to routing all flows on the shortest paths between each origin node and each destination node (see subsection 2.4.1). Typically, a transportation infrastructure does not exhibit this property but rather the level of service it is able to provide is inversely dependent on the amount of mobility demand. At any given time, the higher the demand, the lower the overall system performance will be. In traffic networks such performance degradation takes the form of increased congestion and consequently a dilation of travel time. In this setting, the choice of one traveler influences the choice of others as they all contribute to the increase in congestion. The state where each user is on his/her minimum-cost path given everyone else's choices is referred to as *User Equilibrium* (see section 2.5).

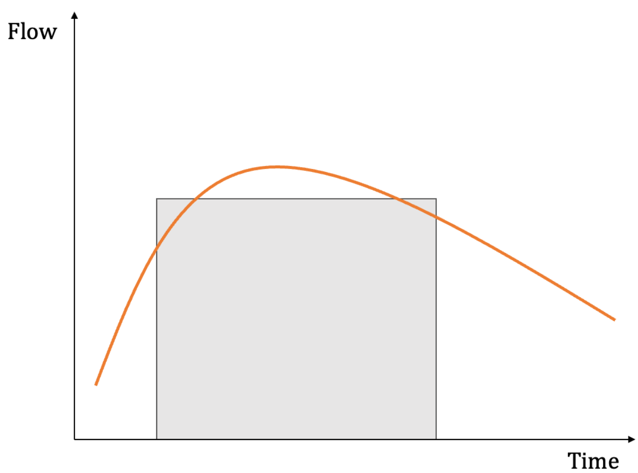



FIGURE 1.2: Static Traffic Assignment Models (STA)

A feasible way, given the demand and functional characteristics of the network, to estimate the flow distribution corresponding to such a state is to solve an appropriate optimization problem (see subsection 2.5.1). On the other hand, when information is imperfect and consequently the users' perception of traffic conditions is affected by an error, the state where each user is on his/her lowest perceived cost path is called *Stochastic User Equilibrium* (see subsection 2.4.4).

As depicted in Fig. 1.2, such approach involves estimating the average traffic flow distribution across the network within a certain period of time. Any flow fluctuationa is omitted from this analysis since no time dimension is taken into account. It is assumed that if both the demand and the infrastructure have been steady for a sufficient amount of time, the actual state of the network at any given time will oscillate somewhere near the equilibrium. For this reason, this class of models is referred to as *Static Traffic Assignment Models* (see chapter 2).

In the context of vulnerability and resilience analyses, static assignment models are typically employed in order to compute the travel time or cost under different scenarios. Since the time dimension is neglected, this approaches consist of comparing the different network equilibria before and after the disruption. The increase in travel time or cost is then utilized as a proxy for the degree of network vulnerability or resilience (Murray-Tuite, 2006; Scott et al., 2006; Chen et al., 2012). For example, Murray-Tuite (2006) employ the DYNASMART-P simulator in order to assess the resilience degree of a network when it is in a user equilibrium or a system optimum state. A system optimum corresponds to a flow pattern where the total cost or travel time of the overall system is minimized (see section 2.5). Chen et al. (2012) rather employ a reliability-based user equilibrium traffic assignment (Shao et al., 2006) for estimating the state of the network. Here as part of the choice process, users also evaluate the reliability level of links. A road with relatively low average travel times but subject to wide fluctuations may be perceived by users as more expensive than another with slightly higher average travel times but with less variability.

Dynamic traffic assignment models (see chapter 3) can be considered a generalization of static models where the time variable and the evolution

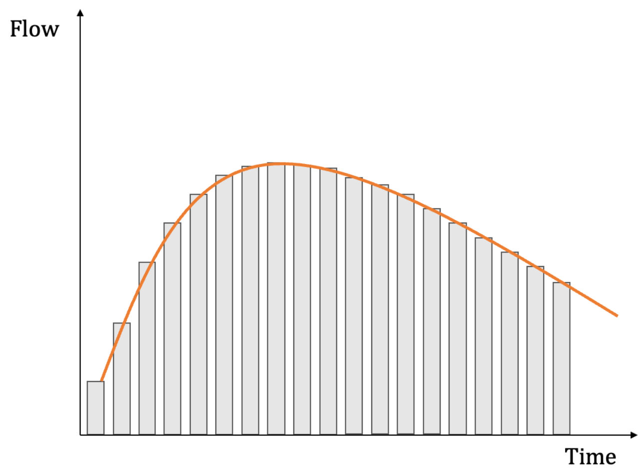



FIGURE 1.3: Fully Dynamic Traffic Assignment Models (DTA)

of quantities related to it are considered. This class of models can further be divided into fully dynamic (or intra-periodic) traffic assignment models (DTA) and semi-dynamic (or inter-periodic) traffic assignment models (DTD).

DTA models are employed to represent network evolution in real time, or at least in the order of minutes (see section 3.1). To this end, those fluctuations that typically occur in traffic flow due to local micro conditions cannot be ignored. For that purpose, DTA models typically consist of two main components namely a user route choice model and a network loading model. Route choice models derived from those employed in static models with the difference being that a choice dimension may be added consisting of the time departure selection and additionally users may change path even during the trip. On the other hand, network loading models make use of some traffic flow model in order to account for local phenomena such as unexpected capacity drops or queue spill-backs. As depicted in Fig. 1.3, DTA models are the most capable at tracking real flow dynamics but at the price of a significantly higher computational effort (Knoop et al., 2012). Therefore, they are usually less frequently employed in vulnerability or resilience analyses when dealing with large-magnitude events whose effects last a long time. On the other hand, they are employed when investigating fluctuations in traffic conditions due to local phenomena with fast dynamics or when the effects of disruption are short-lasting and the system returns to normal operating conditions in a short amount of time (Alam, Habib, and Quigley, 2017; Aslani, Mesgari, and Wiering, 2017; Gauthier, Furno, and El Faouzi, 2018). Sacrificing the scale, this type of analysis allows a significantly finer event representation.

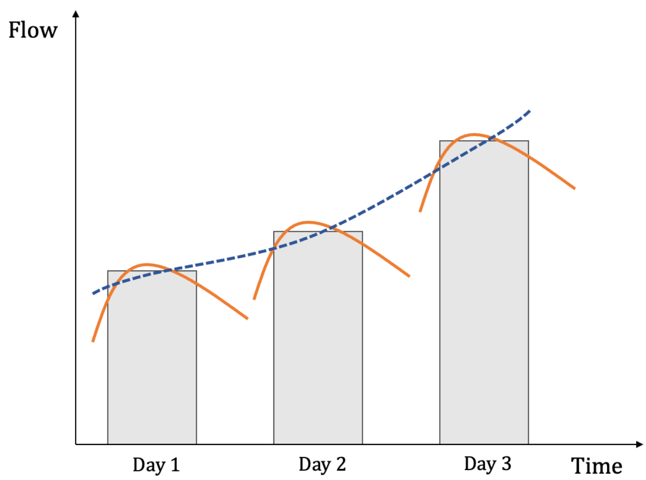



FIGURE 1.4: Semi-dynamic Traffic Assignment Models (DTD)

The class of semi-dynamic models (see section 3.2) lies somewhere in between static and fully dynamic models. The prefix "semi" refers to the fact that these models attempt to capture the equilibration process of the whole system as a series of static assignments associated with subsequent periods. For this reason, they are also referred as inter-periodic or day-to-day traffic assignment models (DTD). When the network due to significant variations in demand or supply falls into a state of disequilibrium, it is reasonable to assume that users will react by possibly changing their choice. For example, if due to a significant change in the infrastructure the travel cost for a certain group of users suddenly increases it is likely that in the future they will shift to routes that now, given the new conditions, have become attractive. Compared with DTA models, the representation of traffic flow dynamics is usually neglected. Rather, the average flow distribution that arises as a reaction to a changing environment is estimated for each period, as depicted in Fig. 1.4. DTD traffic assignment models are considered to be the most appropriate class of models for representing traffic equilibration processes because of their intuitive formulation and ability to accommodate a wide variety of user's choice rules and traffic modeling approaches (Watling and Hazelton, 2003).

Recently DTD models have been employed to assess the long-lasting impact that a disruption have on a traffic network. He and Liu (2012) are probably among the first to employ a DTD assignment model adapted explicitly to analyze the reaction of a traffic network when subjected to a disruptive event. The authors calibrate and validate the model with empirical data collected over 15 days after the collapse of the I-35W bridge in Minneapolis showing a good fit between estimated average daily flows and actual data. Following a similar approach, Wang et al. (2015) propose a day-to-day tolling scheme aiming to increase the rapidity of the system in recovering the the performance lost after a disruption. Finally, Guo and Liu (2011a) investigate the possibility that following significant or long-lasting changes to the network infrastructure, the state of the system may never return to its initial values despite the changes being revoked. For example, the authors hypothesize that following an extended link closure the pattern of flows may permanently change even after reopening. The authors substantiate their argument by showing how after the reopening of the I-35W Bridge, the flows crossing it dropped by 20% despite an unchanged mobility demand. Therefore, this irreversibility phenomenon is modeled within the framework of Bounded Rationality (Mahmassani and Chang, 1987). It is assumed that users, for a variety of reasons related to the complexity of the choice or the pressure of habits, may not exhibit perfect rationality, i.e. their choice strategy may not pursue utility maximization or disutility minimization (see chapter 4). One of the most widely adopted formulations is to consider an indifference band. In other words, travelers are insensitive to excessively minute variations in travel costs and thus unable to distinguish between qualitatively close though not equal alternatives. The resulting implications on the system's steady state are discussed in section 4.2.

## 1.3 Thesis overview and research questions

The objective of this thesis is to develop methodologies designed to estimate the evolution of a transportation network when impacted by a disruptive event which significantly alters its topology. As highlighted in the introduction, purely topological approaches though widely adopted cannot provide a representation for the interaction dynamics emerging between the infrastructure and the mobility demand. Therefore, the main research question of the present thesis can be summarized as follows.

> *How to macroscopically represent the evolution of a transportation network when affected by substantial structural alterations, taking into account the interaction between supply and demand?*

In light of the premises highlighted in the introduction, the approach followed for this thesis is based on dynamic assignment methods, especially inter-periodic ones which provide an appropriate trade-off between the need to represent macroscopically the reaction of travelers on the one hand but without the computational burden associated with more detailed simulation models on the other.

The present thesis is structured as follows. In chapter 2, the state of the art and major contributions to static assignment models theory are presented. Firstly, the traditional framework of transportation analysis and the methods used to represent a transportation network are exposed, with special emphasis on the assumptions regarding users' choice behavior. Subsequently, the literature of deterministic and stochastic static assignment models is reviewed in detail. In section 2.5, the characterization of deterministic equilibrium steady state is discussed in detail and a viable optimization-based methodology for its computation is presented.

In chapter 3 the state of the art regarding dynamic assignment models is reviewed. The main approaches are discussed and contextualized to their respective application domains. Inter-periodic (or day-to-day) assignment models are then explored in more detail and two main frameworks are introduced and discussed.

In chapter 4 an in-depth review of travelers' bounded rationality concept is carried out. Major empirical evidence supporting this idea is reported alongside the main contributions regarding day-to-day assignment models integrating bounded rationality. A formal characterization of the resulting equilibrium configurations is then provided and compared with the counterparts from traditional assignment models.

The second part of the present thesis is devoted to the exposition of three main contributions. The first contribution is presented in chapter 5 where the primary research question is further specified in the following sub-question.

> *How to account for user inertia associated with supply side variation by employing static assignment models within a dynamic framework?*

To this end, a link-based day-to-day traffic assignment model is presented and discussed in detail. The model portrays the flow re-equilibration mechanism of a transportation network affected by a disruptive event. The peculiarity of the approach is that it takes into account user habits that are assumed to affect the evolution of the system following the disturbance. The users' choice selection process is typically represented by multi-path assignment, i.e. the flows from each origin to each destination are distributed over multiple paths depending on traffic conditions. In order to reflect users' inertia towards adaptation, the proposed model constrains a series of successive assignments to specific subsets of paths which over time tend to enlarge if the increase of user travel costs exceeds a certain tolerance threshold.

When a transportation system undergoes a particularly significant disruptive event, a partial modal shift is likely to occur if the transportation system enables it. In this regard, the following research sub-question is formulated.

> *How to take into account the interdependence between different transportation modalities?*

In chapter 6, this shift dynamics is investigated for a multi-modal multi-class transport network. The travel modes considered are road and rail transport. The model represents how passenger flows and freight flows react to disruption according to different behavioral logics. The model includes two stages. In the first one, two traffic assignments are performed for passenger flows and freight flows, where in the first one users aim to minimize their individual travel cost while in the second one freight vehicles move in order to minimize a generalized fleet cost. At the second stage, traveler decisions are implemented and traffic flows loaded onto the network by employing a discrete-time dynamic flow model. The model is then tested in two scenarios prior to and after a disruption. Employing a multi-modal network allows to represent the potential shifts in transportation modality, while adopting a dynamic model allows for more accurate estimation of traffic status-related quantities such as average speeds and pollutant emissions.

Traditionally, users' bounded rationality is represented by means of an indifference band that prevents them from reacting when the stimulus variation is excessively small. In the transportation context, this results in travelers who are insensitive to excessively minute changes in travel costs. In the literature it is also pointed out that users may be affected by "choice supportive bias", i.e. users are inclined to associate positive attributes to the choices they have made while, conversely, they are more likely to associate negative attributes to the alternatives they have not chosen. Within the route choice process, a clear correlation is detected between travel time overestimation and whether a route was chosen or not. Based on these observations, the following sub-questions are formulated.

> *How to account for user inertia associated with routes topology?*
> *Does favoring a route over another partly on the basis of its topology prefigure a bounded rationality route choice process?*

To this end, in chapter 7 a path-based proportional-switch day-to-day assignment model is presented. In the present work, it is suggested that users operate their travel choices influenced not only by the actual path travel costs but also by the topological similarity between them. In other words, if changing traffic conditions on a path leads a user to reconsider his travel choice, he will prefer shifting towards those paths most similar to the one he is abandoning. It is thus shown that such preference implies a bounded rationality behavior. More specifically, a rigorous proof is provided showing that a steady state achieved by the system necessarily corresponds to a Boundedly Rational User Equilibrium (see section 4.2). Furthermore, a method for estimating the relative indifference band is proposed. The model is then applied to two networks affected by a disruption and the results are discussed.

Finally, in chapter 8 some final conclusions and potential research extensions are outlined.

# Part I

# Background

# Chapter 2

# Static Traffic Assignment Models

The study of a traffic network requires the planners to be able to predict the state/s that would most likely occur in case a certain scenario arises. Such a knowledge can be used whenever the infrastructure is involved in significant changes or when it is necessary to predict the users-network interaction when certain conditions vary. In order for any of these analyses to be reliable, it is necessary to approach the problem in the most holistic way possible, once the increase in complexity is taken into account while considering multiple elements within the system description. The amount of traffic flowing at any given time along any street, intersection or square emerges from the simultaneous choices made by a multitude of individuals deciding when to leave, where to go and how to get there. In turn, individual choices are interdependent and the resultant outcome is a trade-off between multiple individual wills grappling with a limited resource. The "finiteness" of the infrastructure resource takes the form of *congestion*. The more users decide to use the infrastructure at the same time the more the system performance will degrade. This effect in turn changes user behavior within a circular dynamics. There is therefore a mutual dependence between user mobility demand on one hand and mobility offer provided by the transportation system on the other (Cascetta, 2013). Together these two components determine the level of congestion, which retroactively influences user choices leading to the actual traffic flow pattern. In the transport engineering field such a flow pattern is denoted as *equilibrium*.

Equilibrium analysis is therefore a viable approach to study transportation networks and to understand their essential characteristics in order to be able to predict, to some extent, their evolution when a variety of factors change. Equilibrium refers to the state at which there are no net forces at work pushing the system towards another state. At the same time, if the system at a certain instant is in a state of disequilibrium, we expect that it will be attracted to the state of equilibrium and will be able to reach it in a certain finite amount of time. While in reality the degrees of freedom available to the user's would prefigure a potentially infinite amount of outcomes, both in the destination location as well as in the ways to get there, several daily's contingencies together with a natural tendency to minimise the travel cost, make the user's behaviour on the network predictable, at least on average, especially when the estimation process covers highly aggregate measures. This allows a number of simplifying assumptions to be introduced during the analysis process, including considering a limited number of origin and destination points or establishing a universal behavioural logic that determines how users will move between these points, being confident of obtaining estimations sufficiently close to the real data.

Defining the most convenient set of assumptions for the case study and computing the resulting traffic equilibrium state falls to *traffic assignment models*. Give, as input the functional characteristics of the network and the transportation demand, a traffic assignment model allows to estimate the traffic flow load on each arc of the network according to a certain mathematical rationale designed to represent the users' behaviour.

## 2.1 Preliminaries

Traffic assignment models constitute the fourth stage within the traditional transportation planning process characterized by the following four steps (Potts and Oliver, 1972; Cascetta, 2013):

1. **Trip Generation:** estimate of the amount of trips, usually expressed in terms of flows, that are generated and/or attracted by specific region of the network. The estimation relates the trip generation to specific variables such as demographics, income, etc. acquired throw surveys.

2. **Trip Distribution:** using the output produced by the previous step, the trips between each origin region and each destination region of the network are estimated. The result is usually expressed in terms of an origin-destination (OD) matrix. Several methodologies have been developed for estimating the mobility demand such as entropy models (Wilson, 1967), gravity models (Voorhees, 2013), the Hitchcock model (Hitchcock, 1941) or the opportunity model (Stouffer, 1940).

3. **Modal Split:** the flows between each origin-destination pair are split between different travel modes such as car, train, bus, bicycle,... based on discrete choice theory (Ben-Akiva and Steven, 1985).

**4. Traffic Assignment:** the flows of each OD couple are loaded on the network according to some mathematical principle designed to represent the route choice behaviour of users and as a consequence the amount of flow on each link is estimated.

The whole four-step process is suitable for long-term planning considering the amount of effort required to deliver it. Nevertheless, it represents an essential base for a series of real-time oriented estimations needed when the system undergoes a highly dynamic phase due to sudden changes in conditions.

Traffic assignment models are extensively used in a wide variety of areas from transportation planning (Overgaard, 1967; Eash et al., 1979; Bliemer et al., 2017; Black, 2018), to transportation network resilience analyses and/or enhancement (Zhang, Mahadevan, and Goebel, 2019; Kaviani, Thompson, and Rajabifard, 2017; Murray-Tuite, 2006), to applications for environmental impact mitigation (Wang et al., 2018), to parking process optimization (Zhang, Mahadevan, and Goebel, 2019; Pel and Chaniotakis, 2017; Chaniotakis and Pel, 2015), to location problem for electric vehicle charging stations (Huang and Kockelman, 2020; Ferro et al., 2021).

The application of assignment models is not limited to road networks, but has its place in a huge range of problems involving other modes of transportation. In the railway field, Lansdowne (1981) proposed an assignment model for routing freight over rail network which is controlled by several carriers. The objective is to minimize the number of interline transfers. Lin, Fang, and Huang (2019) analyzed how to exploit the railway system to its maximum potential by increasing its performance through ticket price adjustments while Wu et al. (2013) formulated a day-to-day assignment model based with boundedly rational travelers. Through relaxing the rational behaviour hypothesis, the authors try to capture correlations over time of users' route choice within the field of railway transport. Xu et al. (2017) proposed a dynamic traffic assignment model for urban rail networks that considers the effects of congestion on queueing dynamics. Moreover, Shengguo and Zhong (2011) illustrated a path flow estimation method that exploits users' entry and exits time records at each station.

Furthermore in shipping, assignment methods have been integrated for example within the berth problem by Venturini et al. (2017). A freight assignment model and a ship routing method are integrated and applied to the Northen Sea Route problem by Lin and Chang (2018). Agrawal and Ziliaskopoulos (2006) developed a dynamic freight assignment model in order to investigate the market equilibrium that emerges from shipper-carrier relationship when no shipper is able to reduce its cost by choosing another carrier. This shipper-carrier relationship is modelled similarly to travellers-road infrastructure interaction.

In the field of air transport, assignment models are generally used within strategies for reducing environmental noise (Ganić et al., 2018; Netjasov, 2008) or lowering pollutants (Mirosavljević, Gvozdenović, and Čokorilo, 2011).

Finally, there are plenty of studies where assignment models are applied to networks integrating two or more transportation modes. Dafermos (1982b) proposed for the first time a general model with the aim of studying the equilibrium of multi-modal transport networks by incorporating the Modal-split, the third phase of the transport planning process, directly into the assignment model. An exhaustive comparative study of existing multi-modal models based on their computational complexity is proposed by Nagurney (1984). In the study by Jourquin and Limbourg (2006), the advantages of using equilibrium analysis on multi-modal networks in the case of freight transport is compared against the application of all-or-nothing assignments. Yamada et al. (2009) designed a bi-level model for the optimal design of a multi-modal freight transportation network. At the lowest level a multi-modal multi-class traffic model is executed while the top level determines the best set of actions in order to maximize the benefit-cost ratio of freight transport. In the work by Bingfeng et al. (2017) the overall impact of dedicated bus lanes on a multi-modal network equilibrium is investigated. In order to account for the road capacity reduction for private vehicles, the impedance of the network arcs is modulated. Moreover, a control model applied to a multi-modal network is defined by Si et al. (2012) with the objective of maximizing a social-cost System Optimum function that takes into account not only the total travel time of the network but also other externalities such as pollutants and greenhouse gases emissions.

## 2.2 Network Representation

Modeling a traffic network first requires partitioning the area of analysis into sub-parts called *zones*. Depending on the model scale, a zone may represent a single neighbourhood as well as an entire city. The study area is then represented through the mathematical concept of a graph, typically oriented. Let $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ be an oriented graph where $\mathcal{N}$ and $\mathcal{A}$ are the set of nodes and the set of links, respectively. Each traffic zone is collapsed into a *centroid* which models all possible origins and/or all possible destinations taking place within the corresponding traffic zone. Each graph node can represent a specific piece of the physical network, such as intersections, squares, or other transportation facilities. However, centroids are those nodes from which or to which traffic flows move. The former are called *source nodes* $\mathcal{R} \subseteq \mathcal{N}$, the latter *destination nodes* $\mathcal{S} \subseteq \mathcal{N}$. It is then possible to define a set of origin-destination (OD) pairs $\mathcal{H} = \mathcal{R} \times \mathcal{S}$.

Any properly defined assignment model needs to take the user choice process into account to some extent. Thus, any macroscopic estimation and prediction of the state of traffic on each of the links in the network must be based on some behavioural principle emerging as a result of a set of assumptions made and a set of functional determinants specific to the network that relate to this behaviour. Although different factors may influence users' route choice decisions, as investigated by Prato, Bekhor, and Pronello (2012), it is also true that the level of congestion, the length of an arc and the time needed to travel it represent the main components in the perceived travel cost. In traffic assignment, the fundamentals component that model these three essential aspects are the *link performance* functions, relating the *travel cost* of a link to the level of congestion on it. They incorporate more or less directly the physical and functional characteristics of the specific road section in a function which is typically positive and strictly increasing with the flow. Therefore, the cost of travelling along the link increases as the congestion on the link increases.

Let $x_a$ be the traffic flow on arc $a \in \mathcal{A}$ and consequently $\mathbf{x} = \{x_a : a \in \mathcal{A}\}$ be the link flow vector. The link cost functions can be decomposed in the following three terms:

$$c_a(\mathbf{x}) = \alpha_1 \cdot t_a(\mathbf{x}) + \alpha_2 \cdot w_a(\mathbf{x}) + \alpha_3 \cdot cm_a(\mathbf{x}) \tag{2.1}$$

where:

- $t_a(\mathbf{x})$ is the functions that relates the congestion on links with the resulting travel time.
- $w_a(\mathbf{x})$ is the function linking the waiting time on the link to the link flows vector.
- $cm_a(\mathbf{x})$ is the monetary cost function a priori defined as a function the link flows vector. It can be further subdivided into the following two components: the toll cost $cm_a^{to}$ and the cost of fuel $cm_a^{fu}(\mathbf{x})$, which may depend on the level of congestion.

$\alpha_i$ coefficients are the function parameters having the double purpose of determining the reciprocal weight of the three cost component and at the same time making them homogenous and therefore comparable. In many formulations, the cost of the link is considered to be the *travel time*, as it is generally the determining component of user behaviour. Several empirical studies seem to suggest that it is the primary factor in flow deterrence, although it is not the only one (Sheffi, 1985). In addition, multiple other factors are strongly correlated with travel time and thus mirror its pattern. In practice, the remaining two components are defined as constants on the basis of average values estimated empirically on a case-by-case basis. Therefore without loss of generality in the following we concentrate on mere travel times.

A preliminary formulation of link performance functions when considering exclusively travel times is as follow:

$$t_a(x_a) = \alpha_a + \beta_a \cdot x_a^{\gamma} \tag{2.2}$$

where $\alpha_a$, $\beta_a$ and $\gamma$ are the model parameters. If the intent is to represent the actual travel times and not just the average, an additional error component $\epsilon_a$ can be added. Setting $\gamma$ implies choosing the model's order while

$\alpha_a$ and $\beta_a$ can be estimated through least squares by using empirical measurements. Another viable approach is to rely on a priori proposed model, for example, one of those proposed in the widely adopted Highway Capacity Manual (HCM, 2000). The following formulation is recommended for highway sections:

$$t_a(x_a) = \frac{L_a}{V_a^{free}} + \delta\left(\frac{L_a}{V_a^{q^{max}}} - \frac{L_a}{V_a^{free}}\right)\left(\frac{x_a}{q_a^{max}}\right)^{\gamma} \tag{2.3}$$

where $L_a$ is the length and $q_a^{max}$ is the capacity of the link $a$, while $V_a^{free}$ and $V_a^{q^{max}}$ are the average speed when the link is unoccupied (free-flow speed) and when the flow is equal to the link capacity, respectively. It is worth noting that (2.2) and (2.3) are equivalent once it is defined that $\alpha_a = \frac{L_a}{V_a^{q^{max}}}$ and $\beta_a = \delta\left(\frac{L_a}{V_a^{q^{max}}} - \frac{L_a}{V_a^{free}}\right)\left(\frac{1}{q_a^{max}}\right)^{\gamma}$. A similar function is suggested by the Federal Highway Administration and reported in (2.4).

$$t_a(x_a) = t_a^0\left[1 + \alpha\left(\frac{x_a}{q_a^{max}}\right)^{\beta}\right] \tag{2.4}$$

In this case $t_a^0$ is the free-flow travel time. It worth noting however that $t_a^0 = \frac{L_a}{V_a^{free}}$. Although identical in principle, the latter two formulations allow to relate the model parameters to quantities that are fairly simple to estimate. For numerous other formulations designed ad-hock for specific scenarios, see HCM, 2000.

## 2.3 Wardrop's Principles

Having introduced the concepts of congestion and link performance functions constitute the fundamental elements of any route decision model, it is then necessary to establish how the user relates to these factors. In other words, it is necessary to define what guides the decisions of the user moving through the traffic network. In the following, the two famous principles by Wardrop (1952) on which almost all travel choice models have traditionally been developed are stated.

**Wardrop's first principle** : Travellers choose the route that minimizes their own travel cost.

**Wardrop's second principle** : Travellers choose the route that contributes, along with the choices of all the others, to minimize the total travel cost.

The first principle states that the ultimate goal of each user on the network is to choose a route that guarantees him/her, given the choices of all other users, a minimum travel cost. Then if all users possess complete information about the state of the network it follows that all "non-minimum travel cost" paths will not be used by any user and at the same time those actually used must share the same minimum travel cost. This is a state where no user has, unilaterally, any incentive to change routes because, from his/her point of view, he/she is already on the shortest path possible. In literature this state is often called "*user optimum*" or "*User Equilibrium*" (UE). The former refers to the fat that the user is maximally satisfied. The latter, on the other hand, emphasises the equilibrium (asymptotically stable) nature of this state. In fact, if for some reasons a portion of the flows deviates from a minimum travel time path to another, the congestion of the latter would increase and so would travel times, while the former see its travel time reduced. Users would then have an incentive to return to their initial route. Due to this, after a certain period of time, a redistribution of flows would bring the system back to its original equilibrium. This is of course provided that users make rational choices, based on the first principle and that they have complete and accurate information, as mentioned previously. By relaxing the assumption of perfect information, we assume that users' estimation regarding the state of the network is inaccurate. The resulting equilibrium state at which each user experiences the lowest possible perceived travel cost is called *Stochastic User Equilibrium* (SUE) (Daganzo and Sheffi, 1977). Either way, the first principle tries to establish the rationale behind the personal and to some extent egoistic choices of users and for this reason it underlies all the models whose aim is to intercept this kind of attitude.

The second principle, on the other hand, depicts a scenario in which the combined choices of users lead to a state where the average travel cost is the minimum possible. In other words, the flow-weighted average of the travel times of all users on the network is minimum. This "*System Optimum*" (SO) state generally does not coincide with the "optimum for users"( i.e. the UE), as already demonstrated in the field of economics by the famous work by Pigou, 1920. This only occurs when link travel costs do not depend on the flows, i.e. when no congestion exists. In order to achieve this state regardless of the type of network, it is therefore necessary to influence users choices in some way. Two are the viable strategies: the first involves forcing users to choose the desired paths. This strategy is mentioned as "*involuntary system optimum*". This approach can be found in the filed of industrial logistics, where a central decision maker chooses which routes to use in order to minimize some kind of global cost function. Other examples can be found in railways and communications network. This strategy is also applied in some traffic management systems during exceptional circumstances (Patriksson, 2015). The second strategy, in the absence of any direct power over users, is "persuasive". By modifying the paths' cost function, e.g. introducing/changing a toll according to a congestion pricing strategy (an extensive review can be found in De Palma and Lindsey (2011)), it is possible to drive the free-will of users in such a way as to favour a more efficient use of the system. The resulting traffic state is usually defined as "*voluntary system optimum*". The state reached is typically not an equilibrium and, when the control action stops, observed flows are likely to return to the user optimum.

## 2.4 Literature Review

The vast literature concerning traffic assignment models can be organized on the basis of several criteria. Typically these are: whether link performance functions depend on flows (congested network) or not (uncongested networks), the set of assumptions regarding user choice behaviour (perfect information, rationality, ...), whether such choice models are applied indiscriminately to all users (single-class models) or different models are applied to different group of users (multi-class models), whether the transport demand is considered fixed (inelastic ) or travel cost dependent (elastic), if one or more travel modes are considered (mono-modal or multi-modal) and, finally and most importantly, how the time variable is accounted within the traffic assignment model (static or dynamic models).

In this chapter, an overview of static assignment models is presented. The most classical deterministic network loading models and their further development for congested networks, i.e. deterministic equilibrium models, are therefore presented. For completeness, stochastic network loading models and the corresponding equilibrium models for congested networks are also provided. Finally, a more in-depth discussion is reserved for the User Equilibrium and System Optimum models.

### 2.4.1 Deterministic Network Loading Models

The first attempts to estimate traffic flows on arterial roads date back to the first half of the last century. At a time of rapid expansion of the road infrastructure, the aim of these early studies was to estimate the amount of flow that would spill over from older arterial roads onto newer high-speed ones. Meanwhile, in the '50s considerable progress was made in defining efficient algorithms for detecting shortest paths within a network (Moore, 1959; Bellman, 1958; Dijkstra et al., 1959). Such techniques allowed minimum routes to be identified within a network with fixed link weights. At the end of the decade, the staff of the Chicago Area Transportation Study and the Armour Research Foundation developed the first fully computer aided traffic assignment based on the algorithm by Moore (1959). This type of assignment was of the *all-or-nothing* type, i.e. the mobility demand of each origin-destination pair is loaded onto the shortest fixed-cost route joining the origin with the destination node (Carroll Jr, 1959). The main advantage of these models lies in obtaining an assignment in a single step, once the minimum cost path is computed, which makes them extremely advantageous from both a computational and a conceptual point of view. The major drawback stems from the fact that the interaction between congestion and network link performance is not taken into account (uncongested network), which significantly compromises the results produced by such methods in most real scenarios.

A generic algorithm for *all-or-nothing* assignments can be condensed as follows (Sheffi, 1985):

**Step 0 (Shortest Paths):** For each origin-destination pair compute the minimum fixed cost path.

**Step 1 (Flow Loading):** Load each origin-destination pair mobility demand onto the found paths.

The three major drawbacks are: (1) not being a robust model. Small variations in the estimated link costs can lead to completely different minimum path sets and thus totally different assignments; (2) the choice between two or more shortest paths is completely arbitrary; (3) the relationship between congestion and network performance is not taken into account and for this reason the estimated flows differ widely from those actually observed. Despite not being used as stand-alone methods, all-or-nothing procedures are an essential component of many traffic assignment algorithms widely adopted by researchers and practitioners alike.

In order to partially overcome the above-mentioned limitations, an attempt was made to take into account the impact of congestion through the definition of link's *performance functions* (refer to Section 2.2 for a more in-depth discussion). Briefly, a performance function relate the congestion of a road section with the associated travel cost (usually in the form of travel time). Making use of such link characterization allow to introduce a series of models that iteratively make use of the network travel times updated against the estimated state at the previous iteration and thus actually take into account the congestion and remaining capacity of the network links. At each iteration an all-or-nothing assignment is performed and the travel times of the arcs are consequently updated using performance functions and used as input for the assignment at the next iteration and so on. This category of assignments is referred to as *capacity-restrained methods*. The quantal loading procedure, developed under the Chicago Area Transportation Study (1960), is an early example of this category of models. The procedure involves zoning the network and then representing it by nodes, each one associated with a specific zone. A source node is then chosen randomly and by an all-or-nothing assignment the demand is routed to the corresponding destination nodes. The travel times are then updated according to the accumulation of flows on the links. Another origin node is therefore selected from the remaining ones and the procedure is then repeated. The methodology differs from a pure all-or-nothing assignment only in the set of travel times used, which in this case evolves at each iteration, and for this reason the computational effort is basically the same. Although quantal loading represents nonetheless an evolution, for each OD pair the demand is still assigned to one path only and, in addition, the estimated flow pattern is highly dependent on the source node sequencing within the procedure.

A possible extension of this model is the one proposed by the Bureau of Public Road reported into the Manual of Traffic Assignment (1964). The steps of the algorithm are as follows: using an all-or-nothing assignment the demand of each OD pair is routed over the network based on the free-flow travel times of the arcs. The congestion level of each arc is then updated and travel times are then recalculated. At the next step the procedure is repeated and the demand of all origin-destination pairs is reassigned to the shortest paths in the network. The procedure stops when a predetermined number of steps has been reached or if the routes selected for each pair do not change between one step and the next. The implicit assumption of this type of procedure is that all users of all origin-destination pairs eventually decide to change route if the shortest one turns out not to be the one used in the previous iteration. The main weakness of this methodology is that no convergence is guaranteed and indeed, depending on the complexity of the network, it is common that after a certain number of steps the estimated flow patterns repeat according to an oscillatory dynamic (Van Vliet, 1973). To overcome this issue, new procedures have been developed where, between one step and the next, only a fixed portion of the users is reassigned. Given an assignment represented by the link flow pattern $\mathbf{x}^n$, at the n-th iteration, once the travel time are updated, a second assignment $\mathbf{y}^n$ is then computed. The network state at iteration $n+1$ is obtained as follows:

$$\mathbf{x}^{n+1} = \mathbf{x}^n(1-\alpha) + \mathbf{y}^n\alpha \tag{2.5}$$

where $\alpha > 0$ represents the portion of users who have changed paths. In "Traffic assignment: methods, applications, products" 1972 the number of steps is fixed a priori at a certain value $K$ and consequently $\alpha = 1/K$. In this case therefore, (2.5) is in fact a smoothing process. Almond (1965) adopted a similar approach but without fixing the number of steps a priori. Several researchers have pointed out that, when the number of steps is sufficiently high, the oscillatory phenomenon reappears. To overcome this problem Smock (1963) proposed that the coefficient $\alpha$ decreases as the number of steps increases, i.e. that $\alpha = 1/k$ at the $k$-th iteration. Almond (1967) suggested a further generalization where (2.5) is replaced with

$$x^{n+1} = \alpha_k x^n + \beta_k y^n \tag{2.6}$$

where $\alpha_k + \beta_k < 1$ for the first iterations. Requiring the sum of the coefficients to be strictly lower than one for the first iterations is meant to implicitly represent the fact that only a fraction of the mobility demand actually enters the network at a given time while the remaining portions enter only at a later time. It should be pointed out that the aim of these methods is not to represent a realistic time-varying dynamics of the evolution of the network, but rather to roughly incorporate some elements of this dynamics, with the aim of obtaining a more realistic final estimation. As pointed out by Patriksson (2015), these first heuristics are remarkably similar to the famous *Frank Wolfe algorithm* by (Frank and Wolfe, 1956) one of the most widely adopted methods for computing equilibrium assignments where the only difference lies in the choice of the step size $\alpha$.

Parallel to the previous heuristics, there is the category of *incremental assignment* methods which attempt to solve the main problem of the previous methods: the demand of each origin-destination pair is routed only on one path. Munby (1968) and Ruiter (1968) approached the problem in the following way: at each iteration only a predetermined fraction of the demand is assigned for all pairs. If the number of iterations and the fraction of demand assigned at each iteration are chosen appropriately, the final state of the network computed approximates an equilibrium much better than any single-route assignment. However, the problem of defining a generalized strategy for determining the fraction and the number of steps remains unsolved. In addition, once a given percentage of demand has been allocated on a route, it can no longer be removed. Vliet (1976) tried to alleviate this shortcoming by removing the worst paths computed by the procedure in favour of a partial flow redistribution to the currently shortest ones.

### 2.4.2 Deterministic Equilibrium Models

*Equilibrium assignments* can be interpreted as a variation of capacity-restricted models. Similarly, they make use of a series of all-or-nothing assignments where at each iteration the link travel times are updated according to the new network configuration. The key difference, however, lies in the choice of algorithm step size. The network states estimated by this category of models generally correspond to one of the two Wardrop (1952) principles. As already mentioned, the first establishes that the users organize themselves on the network in such a way as to minimize their own individual travel cost and this leads to a network state known as User Equilibrium (or User Optimum). The second establishes that the users behave in such a way as to minimize the network's total travel cost resulting in a System Optimum state.

The fundamental work by Beckmann, McGuire, and Winsten (1956) redefines the two principles in terms of convex optimization problems. Wardrop's principles correspond in fact to the optimal solution of a convex optimisation problem with linear flow conservation constraints. At the same time, an iterative algorithm for solving convex quadratic optimization problems is presented (Frank and Wolfe, 1956). This algorithm has been and still is widely used for solving traffic assignment problems in the form of optimization problems. Its great popularity is due to its relatively easy implementation for traffic assignment problem where the algorithm turns out to consist of a series of all-or-nothing assignments coupled with one dimension optimization problems to determine the optimal step size. The calculation of an optimal step, as already mentioned, is in fact the real novelty of this algorithm compared to capacity-restrained heuristics. Briefly, at each iteration for each origin-destination pair, the shortest path is identified and all traffic demand is loaded onto it. The algorithm is then updated again by using a formulation analogous to (2.5), where in this case the new all-or-nothing assignment is represented by $\mathbf{y}^n$. The key difference is that $\alpha$ is determined at each iteration by solving an optimization problem in order to get the best step towards the optimum. Given its centrality within the discipline of traffic assignment, a discussion of UE and SO and the related Beckmann problems is given in section 2.5.

Although these papers were published alongside the development of the heuristics described above, their exploitation for solving traffic assignment problems did not take place until 10 years later and through the work of operational researchers. According to Boyce (1984), the lack of a rigorous scientific approach by transport planners is the reason for the delay between formulation and actual adoption of more sophisticated methods.

These fundamental results have sparked off a prolific line of research. Dafermos and Sparrow (1969) formalized the necessary and sufficient conditions for the existence, uniqueness and stability of the solutions for SO and UE problems. The authors show how, given an SO problem, it is possible to define an associated "UE type" problem and vice versa. This important result allowed to apply the same solution methods to the two classes of problems indiscriminately. Smith (1979) and Dafermos (1980) discussed the existence, uniqueness and stability of solutions of traffic equilibrium problems via variational inequalities. The variational inequalities formulation allows to deal with a wider range of networks characterized by non-separable not symmetrical link cost functions. Non-separable means that the travel time (or travel cost) of a link may also depend on the amount of congestion on other links. Non-symmetrical means that the relationship between one link and another may not be symmetric. A traffic equilibrium model for single-mode and a general multi-modal network is proposed by Dafermos (1982a) while Fisk and Nguyen (1982) analyze and compare the performance of the most common traffic assignment algorithms. Fisk and Boyce (1983) presented an alternative equilibrium-travel choice model which extends the result even when performance functions are not invertible while Nagurney (1984) conducted a performance comparison of the two most widely used methods for solving equilibrium traffic assignment problems for multi-modal networks are compared: relaxation and projection methods. Since both have to solve an optimisation problem at each iteration, the author also compares the performance when adopting the Frank Wolfe and Dafermos-Sparrow algorithms. Hammond (1984) and Marcotte and Guélat (1988) quantify and tested the computational efficiency of various algorithms for solving traffic assignment problems formulated as variational inequalities against network scale and the link performance functions amount of asymmetry. Finally, Gabriel and Bernstein (1997) proposed an assignment model considering also non-additive path travel costs, i.e. with path travel costs not exclusively dependent on the aggregate travel costs of the individual links.

Alongside the research backbone represented by the above-mentioned works, a further one has been explored where other factors potentially affecting the route choice process of users in addition to travel time are taken into account and explored. When users are concerned about arriving on time, the resulting traffic equilibrium is defined as *Risk User Equilibrium*. Each driver considers a safety margin, intended as an extra path travel time, in order to increase the chance of arriving at the destination on time. This concept has been formulated on effective travel time (Hall, 1993) or travel time budget (Lo, Luo, and Siu, 2006) rather than travel times as for the UE. An effective travel time is defined as the sum of the expected travel time plus a safety margin. When on all paths of the same origin-destination pair the effective travel time is the same then the system is at the equilibrium. Uchida and Iida (1993) portray risk users' aversion by defining route cost functionals as the sum of the actual travel time plus the variance of travel times multiplied by a certain coefficient representing users' risk aversion. Moreover, the authors reformulate the SO into the *risk system optimum* mutating the same concept. A similar formulation is proposed by Larsson and Patriksson (1995). Peeta and Ziliaskopoulos (2001) rephrase the same concept of travel time reliability by defining a reliability index $RI_p$ calculated on the basis of the probability that the travel time of route $p$ ends up been greater than a certain tolerance threshold. For a risk-neutral user $RI_p = 0$ for each route, while the coefficient grows as the user's risk aversion increases resulting in an higher perceived path cost. An extension of the aforementioned model in proposed by Nie, Zhang, and Lee (2004) where, in addition to the travel time mean and variance a delay penalty is also considered within users's route choice process.

In addition to the above mentioned studies, the two concepts of *risk adverse user equilibrium* (Bell and Cassir, 2002) and *robust user equilibrium* (Ordóñez and Stier-Moses, 2007) have been also proposed. Despite having quite different formulations, both approaches model users who make their decisions based on worst-case travel times. In the first one, the formulation is based on a Nash game where two different types of players, travelers and demons, compete against each other. Travelers make decisions in order to minimize their travel time while conversely, demons act with the aim of maximizing it. The solution of the game and the related assignment model thus translate into a min max problem. In the second, each user is provided with a travel time budget. For each route, the worst-case travel time is computed by adding up the maximum travel time of each arc composing it. A user does not choose those routes whose worst-case travel time exceeds the budget. A larger budget implies a less risk-averse user.

### 2.4.3 Stochastic Network Loading Models

Originally, the first stochastic network loading models were developed to overcome the intrinsic limitations of all-or-nothing ones. The main criticism leveled at this class of models is that they do not take unpredictable user behaviour into account in any way. In reality, users do not necessarily choose the shortest path when a number of available alternatives have similar travel times/costs and this is even assuming that users make their decisions in order to maximize their utility (or minimize their disutility). The assumption behind stochastic models is that users may not have accurate information of the traffic conditions or sufficient knowledge of the network and therefore their perception of travel times/costs on network links may be imprecise. Despite this key difference, the reasons underlying the user choice process remain similar to those of deterministic models. Again, users are supposed to move across the network in an attempt to minimize their own or collective travel time/cost. The deterministic models presented so far can be rephrased as special instances of the stochastic ones where the probability assigned to the selection of the shortest paths is equal to 1 while the probability assigned to all the others is 0. Stochastic network loading models on the other side assign a non-zero chance of been selected even to suboptimal paths. For this reason the traffic demand of each OD pair spread among multiple paths, generally favoring the shorter ones. For this reason, in contrast to all-or-nothing network loading models, stochastic ones are also referred to as *multi-path models*.

In order to represent the imperfect perception of users, travel times/costs are therefore separated into *objective costs*, which can somehow be measured or estimated by the observer, and a *perception error*, which cannot be directly observed and is therefore modeled by a zero mean random variable. Consequently, average costs are equivalent to objective costs while residuals are independent random variables, identically distributed according to a certain probability function. Stochastic network loading models are usually based on one of the following approaches: *path enumeration* or *path diversion*.

**Path enumeration:** The travel times of links are realizations of random variables. At each iteration, a launch is performed and the travel times of all links are determined. At this point a deterministic assignment model is used. The entire procedure is then executed several times so that link travel times cover sufficiently well the probability distribution of the random variables of which they are realizations.

**Path diversion:** The probability associated with the choice of each route is estimated in advance using some method. The traffic demand is then distributed over the various routes in proportion to the computed probabilities.

Regarding the first category, a first attempt to overcome the limitations of deterministic load models was proposed by Falkenhausen (1968). Assuming that users' perception is imperfect, the link travel times are not defined as constant but are extracted from a log-normal probability distribution. Burrell (1968) fully formalized this idea. The link impedance is assumed to be a random variable, characterized by a certain average i.e. the objective cost, and distributed according to a uniform discrete distribution. The path enumeration is obtained by defining a different set of links for each source node. The ratio between the standard deviation and the actual travel time of each link in the same set is equal. The costs associated with each link are then defined as follows:

$$C = \bar{C}(1 + \gamma \cdot D) \tag{2.7}$$

where $C$ is the random cost of the link, $\bar{C}$ is the actual travel cost, $\gamma$ is the random variable of mean zero and variance unity and $D$ is the diversion factor specific to each source nodes. Mason (1972) reformulate the diversion factor $D$ as the ratio between the standard deviation and the square root of the actual travel cost. In both cases the diversion factor is scaled according to the actual travel cost. As a consequence, between two alternatives, the impact that a given divergence between their respective travel costs has in determining the choice depends on the two actual travel cost. A deviation of e.g. 20% may be insignificant if the actual travel costs are very small to begin with. On the other hand, if the actual travel times on the two paths are substantial, a deviation of the same magnitude may be crucial in shifting the choice in favour of the cheapest one. The stochastic assignment algorithm is characterized by the following steps: (1) at each iteration a source node is selected from those not yet explored; (2) Link costs are randomized using the diversion factor associated to the chosen source node; (3) a minimum-path tree to all other nodes is calculated; (4) flows are routed accordingly. The procedure ends when all nodes have been explored. The aforementioned algorithm can also be used for congested networks, i.e. where the interaction among users and the consequent impact on network performance is not negligible. In addition to this, the same author proposes to include the algorithm within a capacity-restrained procedure where at each iteration only a portion of the demand is actually loaded onto the network and the actual link travel costs are recalculated accordingly to the amount of congestion.

The Dial (1971) algorithm, also known as STOCH algorithm, belongs to the category of path diversion models. It has gained great interest among researchers as it provides a convenient analytical method for estimating the probabilities associated with the use of each path. The objective of the algorithm is to allocate the demand for each origin-destination pair over the possible paths according to certain probability values obtained by replicating a certain logistic distribution. To this end, the residuals of the path cost functions are assumed to be independent and identically distributed Gumbell random variables. The resulting traffic flows attributed to each $k$ route ($f_k$) are then obtained by allocating the OD demand ($d$) in proportion to the estimated probabilities ($P_k$).

$$f_k = d \cdot P_k \tag{2.8}$$

As pointed out by several researchers (Florian and Fox, 1976), Dial's algorithm is in fact a deterministic procedure and the probabilities estimated in practice assume the role of proportion coefficients. More relevant is the problem related to the assumption of independence regarding the variable's random components. The more two paths overlap and thus share a larger portion of the network, the more unreasonable it is to consider them uncorrelated. The resulting effect is that the model overestimates the probability associated with the choice of those strongly overlapping paths (Mayberry, 1973; Daganzo, 1977). Gunnarson (1972) and Tobin (1977) proposed a variant of the STOCH algorithm in an attempt to alleviate its shortcomings. A number of notable variants have been proposed as the C-Logit model (Cascetta et al., 1996), nested logit model (Bekhor, Reznikova, and Toledo, 2007), cross and generalized nested models (Prashker and Bekhor, 1998; Bekhor and Prashker, 2001), the path-size logit model (Ben-Akiva and Bierlaire, 1999; Bovy, Bekhor, and Prato, 2008) as well as the webit-based assignment models (Castillo et al., 2008; Kitthamkesorn and Chen, 2013).

Multi-nominal probit (MNP) is another category of discrete choice models designed to overcome the limitations of logit ones. The random variables are now distributed according to a normal distribution. Daganzo and Sheffi, 1977 formulated a probit route choice model where the perceived travel time of link $a$, named $T_a$, is extracted from a $t_a^0$ mean normal distribution with a variance of $\beta \cdot t_a^0$, thus $T_a \sim N(t_a^0, \beta \cdot t_a^0)$. The coefficient $\beta$ represents the perception uncertainty per unit of travel time. On these premises the travel cost of a route, also a normally distributed random variable, can be derived as the sum of the its link random travel times. Due to the properties of the normal distribution it is possible in this case to establish a correlation between paths sharing one or more links, thus overcoming the limitations of the logit distribution. Unfortunately, it is not possible to determine the probability of choosing a specific route in a closed form. To this end, the authors propose an algorithm incorporating a Monte Carlo simulation. Maher (1992) partially overcome this limitation proposing a stochastic network loading method where discrete choice probabilities are estimated using Clark's approximation (Clark, 1961). This method, although not requiring explicit path enumeration, is computationally onerous and at the same time becomes less precise when the number of available choices (paths) increases (Sheffi, 1985). For these reasons, in a scenario with a real network, it is generally preferable to rely on some sort of simulation.

### 2.4.4 Stochastic Equilibrium Models

Similarly to deterministic models, stochastic network loading models have also been used as a component of various methods to handle congested networks. For the first time Daganzo and Sheffi (1977) proposed the concept of *Stochastic User Equilibrium* (SUE) as the state where each user cannot further unilaterally decrease his/her perceived travel time/cost. The authors show how SUE is a general case of the more classical UE and how it converges to it when the variance of cost random components fades to zero. Sheffi and Powell (1982) proposed a mathematical programming problem involving SUE computation with a path-flow based formulation while Sheffi, Hall, and Daganzo (1982) exploited a time based formulation. In order to compute the SUE state, Fisk (1980) formulated the Method of Successive Averages (MSA). The algorithm can be briefly summarized as follows: at the $n$ iteration, a stochastic network loading assignment is perform based on the current set o travel times $\{t_a^n\}$. This results in finding a target assignment $\{y_a^n\}$. In order to find the next actual flow pattern it is sufficient to set: $x_a^{n+1} = x_a^n + \alpha_n(y_a^n - x_a^n)$ where $a_n$ is the move size at iteration $n$. At the next iteration travel times are updated again and a new target assignment is computed. The process stops when convergence is met. To ensure algorithm convergence, it is necessary that the move size series is not convergent while the squared move size series is convergent. Given these conditions, it is common to set $\alpha_n = 1/n$ (Powell and Sheffi, 1982). In order to compute the assignments, any stochastic network loading model can be used. As noted by Patriksson (2015) this algorithm resembles the ones employed in incremental assignment estimation by Almond (1967).

The stochastic models presented so far assume the imperfect travel cost perception of users as the only source of stochasticity while the actual link travel costs are known by the modeler and constant. Mirchandani and Soroush (1987) presented a stochastic loading model where an additional source of stochasticity is considered besides the one arising from the users' imperfect perception. In this case it is the network itself that is represented as a stochastic system. In other words, the travel times of each link in the network are inherently stochastic, plus the users' random perception of travel times according to a defined probability distribution that differs for different classes of users. Users therefore associate the intrinsic stochasticity of a route with the risk of choosing it and thus they make their choice accordingly. A similar approach is followed by Shao et al. (2006) where the variability of travel times is determined by daily demand fluctuations. Users tend to prefer those routes with less travel time variability i.e. with higher reliability.

Several works have attempted to incorporate the heterogeneity of users into the choice functions. Dial (1997) discussed the mathematical properties and solution algorithm of a bi-criterion stochastic equilibrium traffic assignment where users perceive differently travel times depending on level of information, habit and individual characteristics. A stochastic probit-based model for public transport network is proposed by Nielsen (2000) which takes into account heterogenous choice functions for passengers while Jou (2001) modeled the impact that pre-trip information has in influencing the choice behaviour of different classes of users

Although SUE models overcome the limitations of the overly rigid deterministic counterpart in determining which paths should or should not be used, they simultaneously introduce another problem. A SUE model assigns non-zero probability (and thus an equivalent portion of demand) to all paths regardless of cost. In many choice problems this would not be a problem, unfortunately path choice problems may involve a large number of alternatives. Choosing the right subset of alternatives, or *Master Choice Set*, is not obvious and can significantly influence the outcome of the prediction. The explicit generation of the paths belonging to this set may therefore be computationally expensive. At the same time, however, it allows greater freedom in the definition of path choice behaviour (Bovy, 2009). Several strategies have been employed for the pre-determination of this set. In Leurent (1997), the set is delimited by distance. Only routes that do not exceed a maximum distance should actually be considered while in Leurent (2005) only the paths with a cost not exceeding the absolute minimum are taken into account and, similarly, in De La Barra, Perez, and Anez (1993) a heuristics is proposed to compute a set of k-shortest paths. A probabilistic set generation technique based on the concept of intermediate degrees of availability/perception of each alternative is proposed in Cascetta and Papola, 2001. The utility functions of the paths are augmented by means of a logarithmic component called inclusion function, which is linked to the particular path and thus reshapes its probability of choice. In Bekhor, Toledo, and Prashker (2008) the convergence and processing time of path-based assignment models is analyzed in relation to the route choice model and the the choice set size.

## 2.5 User Equilibrium and System Optimum Traffic Assignment

The User Equilibrium assignment model, alongside with its System Optimum variant, is discussed in detail in this section.

Consider a fully connected oriented graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$ where $\mathcal{N}$ defined the set on nodes while $\mathcal{A}$ the set of links. Considering the subsets of all origin nodes $\mathcal{R} \subseteq \mathcal{N}$ and all destination nodes $\mathcal{S} \subseteq \mathcal{N}$, it is possible to define the mobility demand vector $\mathbf{q} = \{q_h : h \in \mathcal{H}\}$ associated with each origin-destination (OD) pair $h \in \mathcal{H} : \mathcal{R} \times \mathcal{S}$. In other words, it is assumed that $q_h$ represents the total number of trips that in a defined time span take place between the origin node and the destination node of OD pair $h$. The sets of all possible loop-free paths connecting an origin node with a destination node is then defined as $\mathcal{K}$. A path $k \in \mathcal{K}$ is uniquely identified by the set of links it consists of. The link-path incidence matrix $\Delta = \{\delta_{ak} : a \in \mathcal{A}, k \in \mathcal{K}\}$ defines this relationship: $\delta_{ak} = 1$ if link $a$ is part of the path $k$ and $\delta_{ak} = 0$ otherwise. Similarly, let us define the OD pair-path incidence matrix $\Xi = \{\xi_{hk} : k \in \mathcal{K}, h \in \mathcal{H}\}$ where $\xi_{hk} = 1$ if $k \in \mathcal{K}$ is a path connecting OD pair $h \in \mathcal{H}$ and zero otherwise.

Regarding traffic flows, let $\mathbf{x} = \{x_a : a \in \mathcal{A}\}$ and $\mathbf{f} = \{f_k : k \in \mathcal{K}\}$ be the link flow vector and path flow vector respectively while $\mathbf{c} = \{c_a : a \in \mathcal{A}\}$ and $\mathbf{C} = \{C_k : k \in \mathcal{K}\}$ be the link travel time(cost) vector and the path travel time(cost) vector respectively. If only additive path costs are considered, i.e. equal to the sum of the travel costs of their links, the relation between the two vectors can be expressed using the link-path incidence matrix $\mathbf{C} = \Delta'\mathbf{c}$, where $\Delta'$ is the transposed incidence matrix.

A feasible link-path flow pattern $(\mathbf{x}, \mathbf{f})$ is then defined as follows:

$$\mathbf{x} = \Delta\mathbf{f} \tag{2.9}$$

$$\mathbf{q} = \Xi\mathbf{f} \tag{2.10}$$

$$\mathbf{f} \geq 0 \tag{2.11}$$

where (2.9)-(2.11) establish the link-path flow relationship, the travel demand satisfaction and the non negativity of flows. The set of all feasible flow patterns $\Theta$ can therefore be defined as follows:

$$\Theta = \{(\mathbf{x}, \mathbf{f}) : (2.9) - (2.11) \text{ hold}\} \quad (2.12)$$

Let the vector of absolute minimum travel costs be defined as $\boldsymbol{\pi} = \{\pi_h : h \in \mathcal{H}\}$, where $\pi_h$ is the minimum travel cost for the OD pair $h$. Thus, it is possible to define a user equilibrium network state.

**Definition 1.** *A network flow pattern* $(\mathbf{x}, \mathbf{f})$ *is a User Equilibrium if and only if the following holds:*

$$f_k > 0 \implies C_k = \Xi_k' \boldsymbol{\pi} \qquad \forall k \in \mathcal{K} \quad (2.13)$$
$$f_k = 0 \implies C_k \geq \Xi_k' \boldsymbol{\pi} \qquad \forall k \in \mathcal{K} \quad (2.14)$$

where $\Xi_k'$ is the k-th row of the transposed OD pair-path incidence matrix, therefore $\Xi_k' \boldsymbol{\pi} = \pi_h$, if $k$ is a path connecting OD pair $h$. Let us define $f_k^h = f_k : \xi_{hk} = 1$ and $C_k^h = C_k : \xi_{hk} = 1$, then (2.13) and (2.14) can be reformulated as:

$$f_k^h > 0 \implies C_k^h = \pi_h \qquad \forall k \in \mathcal{K}_h, \forall h \in \mathcal{H} \quad (2.15)$$
$$f_k^h = 0 \implies C_k^h \geq \pi_h \qquad \forall k \in \mathcal{K}_h, \forall h \in \mathcal{H} \quad (2.16)$$

where (2.15) implies that if a path $k$ of the pair $h$ is used by users, the travel cost on it must be equal to the minimum possible travel cost for the pair, denoted by $\pi_h$. At the same time, unused paths cannot have a travel cost lower than the absolute minimum as defined in (2.16).

Noting that a user equilibrium must also be feasible, the conditions expressed in (2.13)-(2.14) can be summarize in matrix form as follows:

$$\mathbf{f} \circ (\mathbf{C} - \Xi' \boldsymbol{\pi}) = 0 \quad (2.17)$$
$$\mathbf{C} - \Xi' \boldsymbol{\pi} \geq 0 \quad (2.18)$$
$$(\mathbf{x}, \mathbf{f}) \in \Theta \quad (2.19)$$

where operator $\circ$ is an Hadamard (element-wise) product. The same conditions can be expressed alternatively using the notation introduced in

(2.15)-(2.16):

$$f_k^h(C_k^h - \pi_h) = 0 \qquad \forall k \in \mathcal{K}_h, h \in \mathcal{H} \quad (2.20)$$

$$C_k^h \geq \pi^h \qquad \forall k \in \mathcal{K}_h, h \in \mathcal{H} \quad (2.21)$$

$$(\mathbf{x}, \mathbf{f}) \in \Theta \quad (2.22)$$

where (2.20) states that if the flow on a route is strictly greater than zero, its travel cost must be equal to the minimum. Conversely, if the travel cost is strictly greater than the minimum, the flows on that route must be equal to zero. Condition (2.21) implies that $\pi_h$ must be the minimum travel cost for the OD pair $h$ while (2.22) implies the feasibility of a user equilibrium.

It should be noted that the conditions given in (2.20)-(2.22), or equivalently the definition in (2.15)-(2.16), presuppose users whose behavior follows the Wardrop's first principle. If that holds for all users then the equilibrium is stable. Since users act with the aim of minimizing their travel costs and since at equilibrium they are already on the cheapest possible path, any unilateral deviations from the equilibrium necessarily lead to an increase in travel costs and thus an incentive to move back towards the equilibrium.

**Definition 2.** *A network flow pattern* $(\mathbf{x}, \mathbf{f})$ *is a System Optimum if and only if* $\mathbf{x}'\mathbf{c}$ *is minimized.*

In other words, a flow patter is a System Optimum if it minimizes the total travel cost of the network:

$$\mathbf{x}'\mathbf{c} = \sum_{a \in \mathcal{A}} x_a \cdot c_a(x_a) \quad (2.23)$$

and simultaneously it is feasible, i.e. $(\mathbf{x}, \mathbf{f}) \in \Theta$.

### 2.5.1 UE Mathematical Program

A viable approach which allows finding the link flow pattern $\mathbf{x}^{\mathbf{UE}}$ satisfying all the conditions given in (2.17)-(2.19) or equivalently in (2.20)-(2.22) is to solve the following mathematical program.

**Problem 1.**

$$\min z(\mathbf{x}) = \sum_{a \in \mathcal{A}} \int_0^{x_a} c_a(\omega) d\omega \tag{2.24}$$

*subject to*

$$\mathbf{q} = \Xi \mathbf{f} \tag{2.25}$$

$$\mathbf{f} \geq 0 \tag{2.26}$$

*The definitional constraint*

$$\mathbf{x} = \Delta \mathbf{f} \tag{2.27}$$

The problem requires minimizing a function $z(\mathbf{x})$ which is the sum over all links of the integrals of the respective performance functions $c_a(\cdot)$ of the type described in section 2.2, while conditions (2.24)-(2.27) identify once more the set of feasible flow patterns $\Theta$.

By adopting Lagrange multipliers, problem (2.24)-(2.26) can be reformulated as a problem with only non-negativity constraints. Let the vector of multipliers be defined as $\mathbf{u} = \{u_h : h \in \mathcal{H}\}$, the equivalent optimization problem is defined as follows:

**Problem 2.**

$$\min z[x(\mathbf{f})] + \mathbf{u}'[\mathbf{q} - \Xi \mathbf{f}] \tag{2.28}$$

*subject to*

$$\mathbf{f} \geq 0 \tag{2.29}$$

It should be noted that the Lagrangian in (2.28) is expressed in terms of path flows using the relation defined in (2.27).

### Equivalency Conditions

First-order conditions at a stationary point of the Lagrangian described in (2.28) with nonnegativity constraints (2.29) are given:

$$f_k \frac{\partial L(\mathbf{f}, \mathbf{u})}{\partial f_k} = 0 \qquad \forall k \in \mathcal{K} \quad (2.30)$$

$$\frac{\partial L(\mathbf{f}, \mathbf{u})}{\partial f_k} \geq 0 \qquad \forall k \in \mathcal{K} \quad (2.31)$$

$$\frac{\partial L(\mathbf{f}, \mathbf{u})}{\partial u_h} = 0 \qquad \forall h \in \mathcal{H} \quad (2.32)$$

The condition expressed in (2.32) implies that at the stationary point the flow conservation between the vectors $\mathbf{q}$ and $\mathbf{f}$ must hold, i.e. the demand of each pair $h$ must be satisfied. With regard to the conditions (2.30) and (2.31), let us first compute the term $\partial L(\mathbf{f}, \mathbf{u})/\partial f_k$.

$$\frac{\partial L(\mathbf{f}, \mathbf{u})}{\partial f_k} = \frac{\partial z[x(\mathbf{f})]}{\partial f_k} + \frac{\partial \mathbf{u}'[\mathbf{q} - \Xi \mathbf{f}]}{\partial f_k} \qquad \forall k \in \mathcal{K} \quad (2.33)$$

Let the two addends of the right-hand side of the equation be calculated separately. Starting with $\partial z[x(\mathbf{f})]/\partial f_k$:

$$\frac{\partial z[x(\mathbf{f})]}{\partial f_k} = \sum_{b \in \mathcal{A}} \frac{\partial z[x(\mathbf{f})]}{\partial x_b} \cdot \frac{\partial x_b}{\partial f_k} \qquad \forall k \in \mathcal{K} \quad (2.34)$$

The element on the right-hand side of the equation can be rewrite as follows:

$$\frac{\partial z[x(\mathbf{f})]}{\partial x_b} = \frac{\partial}{\partial x_b} \sum_{a \in \mathcal{A}} \int_0^{x_a} c_a(\omega) d\omega = c_b \qquad \forall b \in \mathcal{A} \quad (2.35)$$

$$\frac{\partial x_b}{\partial f_k} = \frac{\partial \Delta_b \mathbf{f}}{\partial f_k} = \frac{\partial}{\partial f_k} \sum_{l \in \mathcal{K}} \delta_{bl} f_l = \delta_{bk} \qquad \forall k \in \mathcal{K} \quad (2.36)$$

where $\Delta_b$ in (2.61) is the b-th row of the link-path incidence matrix. It is then possible to redefine (2.34) taking into account the results in (2.35) and

(2.61) as follows.

$$\frac{\partial z[x(\mathbf{f})]}{\partial f_k} = \sum_{b \in \mathcal{A}} c_b \cdot \delta_{bk} = \Delta'_k \mathbf{c} = C_k \qquad \forall k \in \mathcal{K} \quad (2.37)$$

where as previously mentioned $C_k$ is the travel cost associated with path $k \in \mathcal{K}$.

Regarding the second addend of the right-hand term of (2.33), it can be firstly noted that:

$$\frac{\partial f_l}{\partial f_k} = \begin{cases} 1 & \text{if } l = k \\ 0 & \text{otherwise} \end{cases} \quad (2.38)$$

consequently:

$$\frac{\partial \mathbf{f}}{\partial f_k} = \begin{bmatrix} \partial f_1 / \partial f_k \\ \vdots \\ \partial f_k / \partial f_k \\ \vdots \\ \partial f_P / \partial f_k \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad (2.39)$$

Taking into account that $\mathbf{q}$ is constant and the multipliers $\mathbf{u}$ do not depend on $\mathbf{f}$, therefore $\partial \mathbf{u}'[\mathbf{q} - \Xi \mathbf{f}]/\partial f_k$ can be computed as follows.

$$\frac{\partial \mathbf{u}'[\mathbf{q} - \Xi \mathbf{f}]}{\partial f_k} = -\mathbf{u}' \Xi \frac{\partial \mathbf{f}}{\partial f_k} = -\mathbf{u}' \Xi^k \quad (2.40)$$

where $\Xi^k$ refers to the k-th column of the OD pair-path incidence matrix. Noticing that $\Xi^k = \Xi'_k$ holds, from (2.40) follows:

$$\frac{\partial \mathbf{u}'[\mathbf{q} - \Xi \mathbf{f}]}{\partial f_k} = -\Xi'_k \mathbf{u} \quad (2.41)$$

Then the partial derivate of the Lagrangian in (2.33) can be rewritten considering (2.37) and (2.40).

$$\frac{\partial L(\mathbf{f},\mathbf{u})}{\partial f_k} = C_k - \Xi_k'\mathbf{u} \tag{2.42}$$

With regard to the partial derivative with respect to multipliers $\partial L(\mathbf{f},\mathbf{u})/\partial u_h$, introduced in (2.32), the calculation is rather straight forward:

$$\frac{\partial L(\mathbf{f},\mathbf{u})}{\partial u_h} = \frac{\partial \mathbf{u}'[\mathbf{q} - \Xi\mathbf{f}]}{\partial u_h} = q_h - \Xi_h\mathbf{f} \qquad \forall h \in \mathcal{H} \tag{2.43}$$

where, as defined previously, $q_h$ represents the mobility demand for the OD pair $h$ while $\Xi_h$ is the $h-th$ row of the OD demand-path incidence matrix.

It is therefore possible to explicitly obtain the first-order conditions stated in (2.30)-(2.32) for Problem 2:

$$\mathbf{f} \circ (\mathbf{C} - \Xi'\mathbf{u}) = 0 \tag{2.44}$$

$$\mathbf{C} - \Xi'\mathbf{u} \geq 0 \tag{2.45}$$

$$\mathbf{q} - \Xi\mathbf{f} = 0 \tag{2.46}$$

where (2.44) and (2.45) are obtained by incorporating (2.42), formulated for all paths, into (2.30) and (2.31) respectively, while (2.46) is obtain exploiting the result from (2.43), formulated for all OD pairs, into (2.32).

According to condition (2.45), for each OD pair the path costs cannot be less than their respective multipliers. In other words, each multiplier represents the minimum possible travel cost for the respective OD pair. Thus, the first-order conditions in (2.44)-(2.45) are equivalent to the User Equilibrium conditions expressed in (2.17)-(2.19), once it is noted that an optimal flows pattern must also be a feasible one and therefore (2.22) holds.

### Uniqueness conditions

To ensure that an optimal solution for the Problem 1 is also unique, it is necessary that the performance functions of the arcs $c_a(x_a)$ are monotonically increasing with respect to the flows $x_a$ and fully separable. In other words, the following two conditions must hold:

$$\frac{dc_a(x_a)}{dx_a} > 0 \qquad \forall a \in \mathcal{A} \tag{2.47}$$

$$\frac{dc_a(x_a)}{dx_b} = 0 \qquad \forall a \neq b \tag{2.48}$$

where (2.47) define the monotonicity of performance functions while (2.48) the full-separability, i.e. a performance function of one link does not depend on the flows of the other links.

The requirement for this condition arises once the Hessian of the function $z(\mathbf{x})$ is calculated using (2.47)-(2.48).

$$H_{z(\mathbf{x})} = \begin{bmatrix} \frac{dc_1(x_1)}{dx_1} & 0 & \dots & 0 \\ 0 & \frac{dc_2(x_a)}{dx_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{dc_A(x_A)}{x_A} \end{bmatrix} \tag{2.49}$$

The Hessian is diagonal with positive diagonal elements, therefore $H_{z(\mathbf{x})} > 0$. Once taken into account that $\Theta$ is a convex set, it can be stated that an optimal solution $\mathbf{x}^*$ for Problem 1 is necessarily unique.

The uniqueness of the solution is guaranteed with respect to link flow patterns $\mathbf{x}$, not with respect to path flow patterns $\mathbf{f}$. The link-path incidence matrix $\Delta$ does not necessarily have full rank and as a consequence having fixed a link flow pattern, the linear system in (2.9) can admit infinite solutions with respect to $\mathbf{f}$.

### 2.5.2 SO Mathematical Program

Similar to the User Equilibrium, one viable approach to determine a System Optimum flow pattern is by solving the following minimization problem.

**Problem 3.**

$$\min \tilde{z}(\mathbf{x}) = \sum_{a \in \mathcal{A}} x_a \cdot c_a(x_a) \tag{2.50}$$

*subject to*

$$\mathbf{q} = \Xi \mathbf{f} \tag{2.51}$$

$$\mathbf{f} \geq 0 \tag{2.52}$$

*The definitional constraint*

$$\mathbf{x} = \Delta \mathbf{f} \tag{2.53}$$

where $\tilde{z}$ in (2.50) is the network total travel cost while (2.51)-(2.53) determine again the feasible set $\Theta$. Making use of Lagrangian multipliers $\tilde{\mathbf{u}} = \{\tilde{u}_h : h \in \mathcal{H}\}$, it is possibile to define the following equivalent minimization problem.

**Problem 4.**

$$\min \tilde{z}[x(\mathbf{f})] + \tilde{\mathbf{u}}'[\mathbf{q} - \Xi \mathbf{f}] \tag{2.54}$$

*subject to*

$$\mathbf{f} \geq 0 \tag{2.55}$$

Following the use of the Langragian, Problem 4 has only non-negativity constraints. Below are the first- and second-order conditions regarding the existence and uniqueness of the solution respectively.

### First Order Conditions

First-order conditions at a stationary point of the Lagrangian described in (2.54) with non-nnegativity constraints (2.55) are given:

$$f_k \frac{\partial L(\mathbf{f}, \tilde{\mathbf{u}})}{\partial f_k} = 0 \qquad \forall k \in \mathcal{K} \quad (2.56)$$

$$\frac{\partial L(\mathbf{f}, \tilde{\mathbf{u}})}{\partial f_k} \geq 0 \qquad \forall k \in \mathcal{K} \quad (2.57)$$

$$\frac{\partial L(\mathbf{f}, \tilde{\mathbf{u}})}{\partial u_h} = 0 \qquad \forall h \in \mathcal{H} \quad (2.58)$$

Firstly, the partial derivative of the Langragian with respect to path flows $\partial L(\mathbf{f}, \tilde{\mathbf{u}})/\partial f_k^h$ is calculated.

$$\frac{\partial L(\mathbf{f}, \tilde{\mathbf{u}})}{\partial f_k} = \frac{\partial \tilde{z}[x(\mathbf{f})]}{\partial f_k} + \frac{\partial \tilde{\mathbf{u}}'[\mathbf{q} - \Xi \mathbf{f}]}{\partial f_k} \qquad \forall k \in \mathcal{K} \quad (2.59)$$

The evaluation of the second addend of the right-hand term is identical to (2.42) except that the multipliers are defined as $\tilde{\mathbf{u}}$ instead of $\mathbf{u}$. With regard to $\partial \tilde{z}[x(\mathbf{f})]/\partial f_k$:

$$\frac{\partial \tilde{z}[x(\mathbf{f})]}{\partial f_k} = \sum_{b \in \mathcal{A}} \frac{\partial z[x\tilde{(}\mathbf{f})]}{x_b} \frac{\partial x_b}{\partial f_k} \qquad \forall k \in \mathcal{K} \quad (2.60)$$

once taken into account (2.61) the following hold.

$$\begin{aligned} \frac{\partial z[x\tilde{(}\mathbf{f})]}{x_b} \frac{\partial x_b}{\partial f_k} &= \frac{\partial z[x\tilde{(}\mathbf{f})]}{x_b} \delta_{bk} = \sum_{b \in \mathcal{A}} \delta_{bk} \frac{\partial}{x_b} \sum_{a \in \mathcal{A}} x_a c_a(x_a) \\ &= \sum_{b \in \mathcal{A}} \delta_{bk} \left[ c_b(x_b) + x_b \frac{dc_b(x_b)}{dx_b} \right] \qquad \forall k \in \mathcal{K} \quad (2.61) \end{aligned}$$

The term within square brackets in (2.61) can be interpreted as marginal total travel cost. Let us define this term as $\tilde{c}_b(x_b)$, therefore it can be stated the following:

$$\frac{\partial \tilde{z}[x(\mathbf{f})]}{\partial f_k} = \sum_{b \in \mathcal{A}} \delta_{bk} \tilde{c}_b(x_b) = \tilde{C}_k \qquad \forall k \in \mathcal{K} \quad (2.62)$$

where $\tilde{C}_k$ can be defined by analogy with $\tilde{c}_k$ as the marginal total travel cost for path $k$.

The computation of the partial derivatives of the Lagrangian with respect to $\tilde{u}_h$, as in condition (2.58), is analogous to (2.43). Therefore, it is possibile to define the first-order conditions for Problem 4:

$$\mathbf{f} \circ (\tilde{\mathbf{C}} - \Xi' \tilde{\mathbf{u}}) = 0 \tag{2.63}$$

$$\tilde{\mathbf{C}} - \Xi' \tilde{\mathbf{u}} \geq 0 \tag{2.64}$$

$$\mathbf{q} - \Xi \mathbf{f} = 0 \tag{2.65}$$

where $\tilde{\mathbf{C}} = \{\tilde{C}_k : k \in \mathcal{K}\}$ is the marginal total travel cost vector. Equations (2.63)-(2.65) state that at the optimum, the total travel marginal cost on each used path for each OD pair must be the minimum, otherwise the path is not used.

The first-order conditions for SO optimization problem are fundamentally equal to those in (2.44)-(2.46) regarding the UE problem, although defined on different performance functions. This fact has an important consequence. It is possible to compute an SO exploiting Problem 1 once link performance functions have been expressed in terms of marginal costs $\tilde{c}_a(x_a)$.

### Second Order Conditions

In order for an optimum of Problem 3 to be unique, it is sufficient to show that the Hessian of $\tilde{z}(x)$, with respect to $\mathbf{x}$, is positive defined given the same conditions as in (2.47)-(2.48). Let us first compute the partial derivatives $\partial \tilde{z}(\mathbf{x}) / \partial x_b$:

$$\frac{\partial \tilde{z}(\mathbf{x})}{\partial x_b} = \frac{\partial}{\partial x_b} \sum_{a \in \mathcal{A}} x_a c_a(x_a) = \sum_{b \in \mathcal{A}} c_b(x_b) + x_b \frac{dc_b(x_b)}{dx_b} \tag{2.66}$$

and

$$\frac{\partial^2 \tilde{z}(\mathbf{x})}{\partial x_b \partial x_a} = \begin{cases} 2\frac{c_a(x_a)}{dx_a} + x_a \frac{d^2 c_a(x_a)}{dx_a^2} & \text{if } b = a \\ 0 & \text{otherwise} \end{cases} \tag{2.67}$$

The Hessian matrix is diagonal, with diagonal elements described by (2.67). If the performance functions are convex and they satisfy the conditions in (2.48) then the diagonal elements are also positive. For this reason we can conclude that $H_{\tilde{z}(\mathbf{x})} > 0$. Once taken into account that $\Theta$ is a convex set, it can be stated that an optimal solution $\mathbf{x}^*$ for Problem 3 is necessarily unique.

## 2.6 Equilibrium assignment and network criticality assessment

As illustrated in chapter 1, the criticality of a transport network in terms of vulnerability, robustness or resilience is generally evaluated following two main approaches: topological-based or system-based methods. Topological methods have the advantage of being computationally highly efficient and scaling well with network size. On the other hand, not taking flow dynamics into account, they do not necessarily reproduce realistic results when applied to transportation networks. System-based criticality approaches incorporate traffic characteristics (such as travel times, generalized travel costs, flows, etc.) within criticality metrics and thus overcome the shortcomings of purely topological analyses. In order to estimate traffic characteristics, assignment models are generally used, which on the one hand guarantee results closer to the real system but on the other hand introduce additional modeling complexity and computational burden. In light of these premises, it is not surprising how a considerable effort has been devoted to reducing the computational burden both for equilibrium computation and for its exploitation for network criticality assessments.

The algorithm by Frank and Wolfe (1956) [FW] has been one of the most widely adopted approaches to solve the deterministic traffic assignment problem. Its strengths are that it requires low memory, no path enumeration is needed, and the implementation procedure is straightforward, characterized by a sequence of all-or-nothing assignments. At the same time, it has a significant drawback. Once in the proximity of the optimum, the algorithm asymptotically converges sub-linearly, since the descent directions tend to become normal to the gradient resulting in a zig-zagging behavior.

To overcome these limitations, several link-based algorithms have been defined to enhance local convergence (Fukushima, 1984; Hearn, Lawphongpanich, and Ventura, 1985; Florian, Guálat, and Spiess, 1987) as well more recent path-based ones (Bar-Gera, 2002; Dial, 2006; Florian, Constantin, and Florian, 2009; Kumar and Peeta, 2010; Gentile, 2014; Galligari and Sciandrone, 2018; Babazadeh et al., 2020). Recent advances in computer science and particularly in primary memory technologies have enabled the emergence of increasingly efficient path-based solutions.

One way to efficiently compute the equilibrium is to decompose the problem into subproblems each one dealing with a single OD pair (Larsson and Patriksson, 1992; Jayakrishnan et al., 1994; Chen, Jayakrishnan, and Tsai, 2002; Florian, Constantin, and Florian, 2009; Kumar and Peeta, 2010; Kumar and Peeta, 2014b; Galligari and Sciandrone, 2018). In each subproblem, flows are shifted from the most expensive routes onto the cheapest ones. Depending on the algorithm employed, the flow transfer may occur between paths with higher-than-average cost to those with lower-than-average cost (Florian, Constantin, and Florian, 2009; Kumar and Peeta, 2010), by comparing at each iteration sequentially the costs of only two paths within the same OD pair (Kumar and Peeta, 2014a) or by selecting the maximum cost path and a lower cost path (Javani and Babazadeh, 2017). The path choice between which flows are transferred determines the algorithm descent direction. Equally important for the overall efficiency is the line search that determines the step-size and therefore the amount of flow transferred at each iteration. Jayakrishnan et al. (1994) and Kumar and Peeta (2010), for example, employed a constant step length while Chen et al. (2013) exploit a self-adaptive strategy based of Armijo-Goldstein condition (Armijo, 1966; Goldstein, 1965) aiming to find acceptable step sizes.

Another viable approach is that followed by origin-based algorithms (Bar-Gera, 2002; Dial, 2006; Bar-Gera, 2010; Gentile, 2014; Zheng, 2015). As the name suggests, in this case the assignment problem is divided into subproblems where the flow of a single origin node is assigned to all its destinations. These algorithms exploit the fact that flows move on bushes rooted in origin nodes each of them constituting an acyclic subnetwork. At each iteration a bush is equilibrated, i.e. finding UE flow pattern on the bush. Flows are transferred from longest paths onto the shortest ones, and without cycles the related computations become highly efficient, enabling their use on large-scale networks. Finally, Gentile (2014) proposed an alternative origin-based algorithm where the assignment problem is partitioned with respect to destination nodes.

Regarding system-based metrics estimation, a traffic assignment model is generally applied iteratively. Some global metric, such as total travel time/cost, is computed for the complete network and then, removing the links one by one (full-scan), for each disrupted networks. Following each link removal, a traffic assignment is computed over again thus obtaining the appropriate metrics. These approaches require at least as many assignments as the number of links forgoing exploring all those scenarios where more than one link may be damaged at any given time. Even a full-scan analysis on a large-scale network by removing every combinations of link pairs appears to be computationally infeasible.

Therefore multiple approaches have been developed in an attempt to mitigate this deficiency, such as approaches based on network partition (Erath et al., 2009; Chen et al., 2012) or sensitivity analysis (Luathep et al., 2011). More specifically, Erath et al. (2009) limits the computational burden by restricting the computation to a subnetwork close to the epicenter of the disruption. Conversely, Luathep et al. (2011) avoid the repeated use of traffic assignments by mutating the idea of sensitivity analysis. The chosen system index is computed by means of a single traffic assignment performed on the complete network. After that, the same index in a disrupted scenario is obtained through the application of a first-order Taylor approximation by means of Clark's method (Clark and Watling, 2002) for sensitivity derivatives calculations.

The set of most vulnerable links in the case of simultaneous failure may not simply be the combination of the most critical links in the case of a single link failure, and the set may not be constituted necessarily by connected links or even in close proximity to each other. For this reason, Wang et al. (2016) propose a mixed-integer non-linear program with equilibrium constraints, aiming to determine the combination of links whose deterioration would induce the most increase in total travel cost in the network. The program is solved applying a piecewise linearization approach and range-reduction technique. Following the idea that a particularly severe disruption is likely to affect a localized portion of the network however wide, Jenelius and Mattsson (2012) propose a grid-based approach where instead of assessing the importance of individual nodes or links, the impact resulting from the disruption of an entire cell (consisting of several links and nodes) is evaluated. More recently, Xu, Chen, and Yang (2017) developed a binary integer bi-level program to estimate the upper and lower bounds of network vulnerability avoiding a full-scan approach. The upper-level subprogram maximizes or minimizes the remaining network throughput with a given number of disrupted links, which corresponds to the upper and lower vulnerability thresholds. The lower-level subprogram verifies the connectivity of OD pair under a network disruption scenario without requiring path enumeration.

Over the past two decades greatly progress has been made on the side of computational efficiency for equilibrium estimation and the use of such techniques for assessing transport network criticality. However, the use of equilibrium assignments remains somewhat problematic when dealing with large-scale transportation networks vulnerability assessments due to full or partial scan of links. One promising direction might be exploiting the vast amount of data now available (Anda, Erath, and Fourie, 2017; Zhu et al., 2018) to bridge the gap between purely topological and system-based analyses by being able to relate the topology of a network to a prediction, even an approximate one, of user response under different scenario.

# Chapter 3

# Dynamic Traffic Assignment Models

Dynamic Traffic Assignment Models can be considered a generalization of the static models discussed in chapter 2. Static traffic assignment models, especially equilibrium-based ones, aim to find stationary flow configurations where demand flows, path flows and link flows are distributed on the network congruently to link (path) costs. They consist of finding a flow vector that, given a demand model and a supply model, reproduces itself. On the other hand, dynamic models are not limited to the estimation of network equilibrium configurations but rather they are able to intercept changes in flows distribution within the period of study, whether when they are caused by exogenous or endogenous factors. This ability makes their deployment necessary when the aim is to investigate the time-varying dynamics that emerges from demand-supply interaction.

There are many dimensions over which different subcategories of dynamic assignment models can be identified. The most relevant are:

**Choice Dimension** Assignment models can be divided into those where the path is the only choice available to users (Ben-Akiva, De Palma, and Isam, 1991; Lo and Szeto, 2002; Pel, Bliemer, and Hoogendoorn, 2009) or, alternatively, into those where user's choice regards exclusively the departure time (Vickrey, 1969; Small, 1982). In addition, there are models where users make joint travel choices across both domains (Mahmassani and Herman, 1984; Szeto and Lo, 2004).

**Nature of the time variable** Depending on how time is represented, a distinction can be made between continuous (e.g. Lam and Huang, 1995) and discrete models (e.g. Szeto and Lo, 2006).

**Demand elasticity** Depending on whether or not the transport demand is globally affected by changes coming from the supply side, models can be divided between those with fixed or inelastic demand (Szeto and Lo, 2004) and those with variable or elastic demand (Tong and Wong, 2000).

**Scale of the time variable** Depending on the granularity of the time variable, models can be distinguished between intra-periodic or fully dynamic (Szeto and Lo, 2004) and inter-periodic or semi-dynamic

(Cascetta, 1989) models. For the former, the time unit generally ranges from a few seconds to a handful of minutes, while for the latter, even an entire day.

Among the above classifications, the most important is probably the last one, i. e. the "scale" of temporal variable representation and it will therefore be further detailed.

When the phenomena under study are characterized by fast and localized dynamics (e.g. queue spill back, rapid speed variations), *fully-dynamic traffic assignment models* (DTA) are typically applied. They are referred to as "fully dynamic" because they make use of both a travel choice model and a network loading model which directly represents traffic flow moving on the network. Given the time scale of the representation, ranging from a few seconds to a few minutes, they are also referred to as *intra-periodic* or *within-day* dynamic traffic assignments denoting the representation of the demand-supply interaction occurring on an instant by instant basis.

On the other hand, when an event produces macroscopic, global and long-lasting effects on the network, *semi-dynamic models* are generally used, allowing to study the long-term average consequences of the events. The prefix "semi" refers to the fact that these models attempt to capture the global evolution of traffic patterns as a series of static assignments associated with subsequent periods. For this reason, they are also referred as *inter-periodic* or *day-to-day traffic assignment models* (DTD). Here it is assumed that due to significant changes in the demand or in the supply, the system may be in different feasible states where flow patterns are not necessarily congruent with travel costs (in the sense used for static models).

This chapter is devoted to outlining the main literature pertaining dynamic traffic assignment models. First an overview of DTA models is provided in section 3.1 and then a focus on DTD models is outlined in section 3.2.

## 3.1 DTA Models

DTA assignment models are designed to track users' reactions to real-time changes in traffic conditions. To this end, it is essential to provide a convincing representation of flow propagation over the network. These models are therefore composed of the following two fundamental components: a *travel choice model* and a *network loading model*. Travel choice models, as mentioned earlier, may include a time departure choice models or a route choice models or both. Each decision depends on time-varying quantities such as the OD matrix and the current flow distribution whose estimation is delegated to the network loading model. The modeling approaches and network loading models most commonly used in the definition of DTA are presented in the following.

In order to represent users' travel choices, two approaches are commonly used within DTA models (Szeto and Wong, 2012): *analytical approaches* or *simulative approaches*. The extension of static models into dynamic ones implies, with reference to a particular flow, the need to know when it will leave its origin and when it will be on a specific link. The addition of the temporal dimension greatly complicates the mathematical formulation of any attempt to extend static models while at the same time leading to the non-convexity of related optimization problems (Rakha and Tawfik, 2019). For this reason, scientific research covering DTA models has followed two different approaches. On the one hand, analytical approaches attempt to retain most of the features from static models, such as being able to prove the existence and uniqueness of solutions while making compromises regarding the representation of traffic phenomena and user behavior. On the other hand, simulative approaches move in the opposite direction by characterizing traffic phenomena more precisely but at the same time forgoing any closed-form resolution of the problem or establishing any theoretical properties. In this case, all traffic phenomena and the resulting constraints such as the link-path incidence, vehicle movement, and flow conservations are obtained by simulation.

Analytical approaches extend the concepts developed for static assignments into a time-varying framework. An early attempt to extend the SO formulation to a dynamic context was proposed by Merchant and Nemhauser (1978). The model is formulated as a discrete-time mathematical programming problem limited to the fixed demand and single destination case. The model makes use of static performance functions in order to extract travel costs from traffic flow patterns. Birge and Ho (1993) extended Merchant and Nemhauser's model by developing a non-linear, non-convex stochastic multi-level mathematical programming problem. Friesz et al. (1989) proposed an optimization approach inspired by the objective functions from the Beckmann UE mathematical formulation in order the compute a Dynamic User Equilibrium (DUE). Similarly, Ran, Boyce, and LeBlanc (1993) proposed an optimal control strategy for DUE computation making use of link inflows and outflows as control variables. Alternatively, some formulations with variational inequalities have been proposed for the DUE computation. Friesz et al. (1993) proposed a variational inequalities formulation where both departure time and route choice are taken into account alongside with penalty function related to early or late arrivals. In order to estimate traffic related quantities static performance function are used. To improve the representation of traffic phenomena, Ran and Boyce (1996) proposed a variational inequality-based model that accounts for queue formation. However, the representation of queue dynamics requires adding link's capacity constraints which as a result impact negatively on the computation feasibility. Ziliaskopoulos and Wardell (2000) proposed a linear model for the SO case coupled with a Cell Transmission Model (CTM) (Daganzo, 1994) as a dynamic traffic model for representing flow propagation, overcoming the limits of static performance functions. More recently, Friesz et al. (2011) proposed variational inequalities-based model that exploits a system of ordinary differential equations to represent flow propagation.

Simulative approaches forgo the definition of model mathematical properties in favor of an easier implementation of flow propagation models. Model constraints such as flow conservation and link-path incidence are obtained implicitly at each iteration of the simulation. An early example of this type of approach is represented by SATURN (Simulation and Assignment in Urban Road Networks) which makes use of a series of all-or-nothing assignments within an iterative procedure (Van Vliet, 1982). CONTRAM (CONtinuous TRaffic Assignment Model) (Leonard, Gower, and Taylor, 1989), similarly to SATURN, is made of two components, namely a traffic dynamic model and a traffic assignment model. CONTRAM groups vehicles into "packets" and the routing logic is defined at the packet level and thus affects all vehicles in the packet. Therefore, all vehicles within the same packet are routed along the same minimum cost route, computed from a weighted average of past path costs. An evolution of CONTRAM is represented by DYNASMART (Abdelfatah and Mahmassani, 2001) and INTEGRATION (Van Aerde et al., 2002). Within the same packet, each vehicle is represented microscopically by integrating different microscopic traffic models (car following and lane changing) while there are a variety of assignment models available including Time-dependent Frank-Wolfe Algorithm for UE computation or Time-dependet MSA for SUE computation as well as a more rudimentary distance-based routing. Ben-Akiva et al. (1998) propose DynaMIT as a dynamic traffic assignment scheme designed to estimate current and future traffic conditions in real time. It consists of a demand estimation module and a supply estimation module that interact to generate route directions to induce a UE on the network within a rolling horizon framework. Simulation models presented so far are are considered as mesoscopic since traffic flow representation is handled in an aggregate way at a high level, while the behavioral rules are modeled at a finer scale. Finally, more recent microscopic simulation models are for example AIMSUN (Casas et al., 2010) or VISSIM (Fellendorf and Vortisch, 2010).

Network loading models represent how traffic flows propagate through the network and thus impact the evolution of all traffic-related quantities. Moreover, for all analytical DUE models, network loading models are exploited to compute the path delay operator (Friesz et al., 1993) whose function is to bind to each departure time a vector of travel times, one for each path. Depending on the level of detail of the flow representation, network loading models are typically divided into macroscopic, mesoscopic and miscroscopic models (Hoogendoorn and Bovy, 2001; Rakha, Tawfik, and Meyers, 2009; Ferrara, Sacone, and Siri, 2018). The former represent traffic dynamics in an aggregate way where vehicular flow is represented as a stream, in analogy to a liquid or gas flow, and where equally aggregate quantities such as density, average velocity and hourly flow are defined. The latter, on the other hand, represent traffic at individual vehicle level and explicitly model the interaction among them. macroscopic traffic phenomena are not directly represented but emerge spontaneously from the interaction among a large number of vehicles on the network. Mesoscopic models lie in between and represent some aspects in detail (such as user behavior) and others in an aggregate way (such as vehicular dynamics). Regarding macroscopic models, early attempts were devised to represent vehicular dynamics on a single link through the definition of an exit-function-based model (Merchant and Nemhauser, 1978) and performance-function based models (Janson, 1991). However, these models were unable to represent some traffic phenomena such as queue spillback. The first example of a DTA model incorporating a first-order model based on kinematic-wave (Lighthill and Whitham, 1955) (LWR) was first proposed by Kuwahara and Akamatsu (2001) which was followed by numerous other works (Lo and Szeto, 2002; Friesz et al., 2013; Long et al., 2013). LWR has been resolved in the solution scheme by Daganzo (1994) (Cell Transmission Model (CTM)), by Newell (1993) (Link Transmission Model (LTM)) and more recently within the solution procedure proposed by Gentile et al. (2010) (General Link Transmission Model (GLTM)). First-order models, although extensively used within DTAs due to their simpler mathematical formulation, are unable to represent more complex traffic phenomena such as start-stop waves, capacity drops and internal traffic flow oscillations. In order to overcome these limitations, second-order macroscopic models have been formulated such as the Payne-Withman model (PW) (Whitham, 1974; Payne, 1979) and its evolutions (Aw and Rascle, 2000; Zhang, 2002). First-order models associate one flow value for a given particular density condition on a link one and only one flow and therefore one constant average velocity over the entire link. Second-order models also take into account acceleration phenomena and consequently allow the representation of more complex traffic waves. The first discrete version of the PW was proposed in the late ’80s (Papageorgiou, Blosseville, and Hadj-Salem, 1990) and applied to Boulevard Périphérique in Paris. The PW model is discretized in both space and time and modified in order to take into account the effect that inflow and outflow ramps may have on the mainstream. The model was then extended for the highway context and named METANET (Kotsialos et al., 2002). Alongside the link-based models mentioned above, where dynamics are expressed through quantities related to network links, several macroscopic node-based models have also been proposed where flow dynamics are handled at intersections (Tampère et al., 2011; Gibb, 2011; Corthout et al., 2012).

Compared to Macroscopic models, an opposite approach is followed by microscopic models where instead vehicles are represented individually, often with individual features, and where time is represented at a finer scale and they generally incorporate models to represent users reaction to localized near conditions. Such behaviors are taken into account in microscopic models such as *car-following models* (Reuschel, 1950) like the safety-distance models (Gipps, 1981), the psycho physical models (Evans and Rothery, 1973) and driving errors models (Hamdar and Mahmassani, 2008). Parallel to these, another main class of microscopic models is represented by the *lane changing models* which deal instead with the lateral movements of vehicles between lanes. Lane change dynamics is driven by three main factors: underline motivation, desired lane selection and gap acceptance decision. Gipps (1986) proposes a model in which users’ actions are governed by two drives: the desire to maintain a certain speed as long as possible and the need to be in the appropriate lane before making a turn. A similar approach is proposed for highway context by Yang and Koutsopoulos (1996) where the reasons behind lane changes are separated into external conditions (e.g. a narrowing) or internal conditions (e.g the user’s own desires). Finally, Ahmed (1999) propose a gap acceptance model where the action of changing lanes is a function of the space between two subsequent vehicles occupying the targeted lane at a given time. Another category of microscopic models is that of *cellular automata models* (Nagel, 1998) where the network is represented through a grid of cells within discrete-time dynamics. Generally, a cell can accommodate a vehicle and the dynamics of a road section is characterized by a sequence of states in which each element of the cell grid describing the road topology can be occupied or free. Fundamental traffic quantities such as speed and density are expressed as a function of the cells (Nagel and Schreckenberg, 1992). Microscopic models are often provided as components within traffic simulation software such as AIMSUN (Casas et al., 2010), SUMO (Rakha, Tawfik, and Meyers, 2009) or VISSIM (Fellendorf and Vortisch, 2010). Finally, mesoscopic traffic models fall somewhere in between the above two categories by making use of detailed behavioral patterns but at the same time representing traffic in an aggregate manner. Time headway models (Branston, 1976), cluster models (Mahnke, Kaupužs, and Lubashevsky, 2005) and kinetic gas models (Paveri-Fontana, 1975) belong to this category.

## 3.2 DTD Models

Day-to-day assignment models are intended to represent the aggregate equilibration dynamics of traffic flows that occur on a network when the system happens to be in a state of disequilibrium at a certain moment in time, i.e., when the pattern of flows is not consistent with the pattern of costs given a user behavioral model. The focus, compared to DTA models, is not on vehicle dynamics but rather on the evolution from one period to the next of users' choices in response to a changing environment that influences them and it is influenced by them in a circular process that gradually evolves from a situation of disequilibrium to one of equilibrium. It is assumed that at the beginning of each period the users have the possibility to acquire awareness about the characteristics of the network up to that moment and can therefore act accordingly.

As pointed out by Cantarella and Watling (2016), since his seminal work Wardrop (1952) justifies the proposed concept of equilibrium by alluding to the role that a dynamic equilibration process would play in such circumstances by stating that:

> *"It may be assumed that traffic will tend to settle down into an equilibrium situation."*

The same assumption underlies the following work by Beckmann, McGuire, and Winsten (1956) while it is only much later that Smith (1979) delineates the characteristics of virtually all subsequent DTD models:

> *"Consider a single driver who has travelled at least once today. He may use the same routes tomorrow. However, if he does change a route then he must change to a route which today was cheaper than the one he actually used today."*

The nature of the time variable assumes significantly different connotations with respect of its counterpart in DTA models. In fact, taking discrete models as an example, regardless of the discretization interval in DTA models two successive instants of time are also contiguous. In other words, defined $T$ as the discretization interval and $k$ as the index of the time variable $t$, then $t_k = kT$ holds and consequently $t_k - t_{k-1} = T$. In DTAs context, stating that two instants of time are consecutive implies that they are also "close" from a temporal perspective. By contrast, in DTD domain this relationship is not necessarily valid. As stated by Cascetta (1989), referring to a period as an "epoch":

> *"... epochs can have either a 'chronological' interpretation as successive reference periods of similar characteristics (e.g. the a.m. peak period of successive working days) or they can be defined as 'fictitious' moment in which users acquire awareness of path attributes and make their decision."*

Therefore, two consecutive periods do not necessarily represent contiguous moments and a significant amount of time may elapse between them. Given these premises, it is common to consider time in a discrete way. However, there is no shortage of “exceptions” within this line of research.

The early seminal works in DTD by Horowitz (1984) and Smith (1984) aim to study the stability of equilibrium states as defined by static assignment models. More specifically, Horowitz (1984) analyzed the stability of the SUE on a network of two links by defining a day-to-day dynamics based on the theory of nonlinear discrete-time dynamical systems. Smith, 1984 proposed a model whereby users "shift" from a path to cheaper paths that are available to the same origin-destination pair. The rate of this shift is proportional to the cost difference between the paths at a particular time. This flow swapping process is known in the literature as a *proportial-switch adjustment process* and its major strength lies in the fact that it allows an endless number of different user shift behaviors to be implemented with relative ease. Starting with these early works, research in DTD has developed considerably in the following years into a solid body of literature.

Different classifications of DTDs can be made. First of all, DTDs can be divided into continuous-time models and discrete-time models. Continuous-time models (Smith, 1984; Friesz et al., 1994; Smith and Wisten, 1995; Zhang and Nagurney, 1996; Bar-Gera, 2005; He, Guo, and Liu, 2010; Guo and Liu, 2011b; He and Peeta, 2016) are defined by differential equations and they are usually justified because of their convenience in obtaining analytical or theoretical results. Continuous dynamics can be considered as an approximation of the corresponding discrete models, and the accuracy of this approximation increases if the considered periods are shorter compared to the overall scale of the model. Nevertheless the peculiarities of continuous and discreet models should not be underestimated being that what is valid for the former might not be for the latter (e.g. convergence). Discrete time models (Horowitz, 1984; Cantarella and Cascetta, 1995; Watling, 1999; Zhang, Nagurney, and Wu, 2001; Bar-Gera, 2005; He and Liu, 2012) on the contrary generate a sequence of snapshots of the network, one for each period, in order to capture its evolution. Depending on the implemented route choice behavior model, at the beginning of every period each user makes use of the available information and formulates a decision. The result of the combined choices of all users characterizes the flow pattern of a specific period.

It is possible to make a distinction between deterministic or stochastic DTD models, depending on whether the day-to-day process establishes a one-to-one or one-to-many relationship between the (link/path) flow vectors of two consecutive periods. Stochastic models (Cascetta, 1989; Davis and Nihan, 1993; Cantarella and Cascetta, 1995; Watling and Hazelton, 2003; Hazelton and Watling, 2004; Watling and Cantarella, 2015; Watling and Hazelton, 2018; Hazelton, 2022) aim to intercept the uncertainty and the variability that characterizes real traffic phenomenon by usually represents the evolution of the fundamental system variables via Markovian processes. By contrast, deterministic models assign a unique (link/path) flow pattern to each period. Another important distinction needs to be made regarding deterministic models: DTD models encapsulating deterministic path choice models (Smith, 1984; He, Guo, and Liu, 2010; Smith and Mounce, 2011; Guo and Liu, 2011b; Kumar and Peeta, 2015a; Zhou et al., 2017) and DTD models based on stochastic path choice models (Horowitz, 1984; Cascetta et al., 1996; Watling, 1999; Cantarella and Watling, 2016; Smith and Watling, 2016; Ye, 2022). The former models assume that users have perfect information at their disposal, i.e. at any given time, they know all the relevant network details (usually travel times) needed to formulate their mobility decisions. If the hypotheses about the users' behavior prefigure a rational path choice process as defined by Yang and Zhang (2009), then the stationary state of the system corresponds to a Wardrop's equilibrium. The latter, on the contrary, assume that users lack perfect information and, therefore, their assessments of the network state are affected by an estimation error. As a result, perceived travel times, which vary across individuals, differ from actual travel times. The residuals are modeled by means of random variables given a certain distribution. Different assumptions about residuals distribution lead to different models, among which the most famous one is certainly the Multinomial-Logit model (Watling, 1999; Guo, Yang, and Huang, 2013; Smith and Watling, 2016) allowing to find the path choice probability solution in closed form. Other notable variants have been proposed, such as the C-Logit model (Cascetta et al., 1996), the path-size Logit model and the nested Logit model (Ben-Akiva and Bierlaire, 2003; Yu, Han, and Ochieng, 2020), to alleviate some of the weaknesses of Logit-base models, as well as more recent day-to-day weibit-based models (Castillo et al., 2008; Kitthamkesorn and Chen, 2013; Xu et al., 2021). The stationary state of these models leads to a SUE. Note that, despite incorporating a stochastic route choice behavior, these models are fully deterministic, i.e any given pattern of costs is associated with one and only one vector of path choice probabilities. The total flows at the next period are then distributed proportionally to the computed probabilities.

Given the initial conditions, the trajectory of the system is unique.

Moreover, as argued by Tan, Yang, and Guo (2015), day-to-day user path choice behavior has been declined into five major categories of dynamic processes: simplex gravity flow dynamics (Smith, 1983), network tatonnement process (Friesz et al., 1994; Jin, 2007; Guo and Huang, 2009), projected dynamical system (Zhang and Nagurney, 1996; Nagurney and Zhang, 1997), evolutionary traffic dynamics (Sandholm, 2001; Yang, 2005) and the aforementioned proportional-switch adjustment process (Smith, 1984; Smith and Wisten, 1995; Huang and Lam, 2002; Peeta and Yang, 2003; Mounce, 2006; Mounce and Carey, 2011). In particular, the proportional-switch adjustment process has been quite popular due to its simple formulation and intuitive behavioral interpretation. For instance, Cho and Hwang (2005) define a variation of the original Smiths model based on a stimulus-reaction dynamic whereby the diverting flow is proportional not only to travel times differences but also to users' sensitivity to them. Smith and Mounce (2011) formulated an adjustment process based on split rates at nodes in order to mitigate some of the problems associated with deterministic path-based models as noted by He, Guo, and Liu (2010). Li, Tan, and Chen (2012) designed an excess cost dynamic where path choice behavior depends on the difference between the costs experienced by users and a certain reference value. A non-linear pairwise proportional-switch-based model is defined in Zhang et al., 2015a. Finally, the proportional-switch mechanism has been incorporated into a mixed equilibrium model (Zhou et al., 2017; Wang, Peeta, and He, 2019) for assessing the effects that ATIS-induced behavior and autonomous vehicle penetration have on the network respectively as well as part of a dynamic assignment model which takes into account the presence of electric vehicles (Agrawal et al., 2016).

It is worth noting that the progressive adjustment mechanism of vehicular flows in response to demand-side or supply-side fluctuations is well suited to being represented by a *path-flow dynamics* in which the state space is expressed by means of path flows. Nevertheless, starting from the work by He, Guo, and Liu (2010), a second stream of DTD models has been developed where the dynamics are described by means of link flow. This category of models is referred to as *link-based DTD models* (Guo and Liu, 2011b; Han and Du, 2012; He and Liu, 2012; Guo, Yang, and Huang, 2013; Guo et al., 2015; He and Peeta, 2016; Siri, Siri, and Sacone, 2020b). The evolution of the network is usually handled by a series of periodically computed traffic assignments under changing conditions. As pointed out by He, Guo, and Liu, 2010, not having to rely on path flows avoids a substantial problem that afflicts path-based models (indeed only deterministic ones): the need to estimate an initial path flow pattern that, given the corresponding link flow pattern, is generally not unique and therefore not uniquely identifiable. The initial state of the system in fact is reasonably estimated using classical static assignments which, in the deterministic case, are known to generate unique solutions in the domain of link flows but not in the domain of path flows. For the related mathematical discussion, the reader is referred to section 2.5. The essential problem is that for path-based dynamics, different initial path-flow patterns result in significantly different path-flow trajectories. Aside from that, link-based DTD models also have their shortcomings. They do not allow to estimate "how far" from a congested (or disturbed) area flows start shifting in order to avoid that area. This aspect limits the range of control strategies that can be used coupled with link-based DTD models. But more importantly, link-based DTD models do not allow the adoption of path choice models which have straightforward behavioral interpretation (Kumar and Peeta, 2015a). For these reasons, several techniques have been developed to estimate, out of a link flow pattern, the most reasonable corresponding path flow pattern and thus making the use of a path-based model plausible in these cases as well (Rossi, McNeil, and Hendrickson, 1989; Larsson et al., 2001; Bar-Gera, 2005).

## 3.3 General Frameworks for DTD Assignment Models

In this section, two general formulations for link-based and path-based DTD assignment models proposed by Guo, Yang, and Huang (2013) and Guo and Huang (2016) respectively are presented.

### Link-based formulation

Since the work by He, Guo, and Liu (2010), there has been a growing interest in defining DTD models having a dynamics defined on the link flow patterns instead of the more used path flow patterns. This approach was justified from some observations made by the authors regarding two deficiencies inherent in DTD path-based models. The first one is that a DTD path-based model's trajectory changes significantly by starting from different initial flow patterns. This is a significant problem because different path flow patterns imply different corresponding link flow patterns on which network performance metrics are generally defined. The second issue is that DTD path-based models typically neglect the interdependence between paths.

Following the work by He, Guo, and Liu (2010), in which a continuous time link-based DTD model is proposed, Guo, Yang, and Huang (2013) suggested a general framework for discrete time link-based DTD models. This section is devoted to briefly introducing this model, adopting whenever possible the notation presented in section 2.5.

Consider a traffic network represented by an oriented graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$ where $\mathcal{N}$ is the set of nodes and $\mathcal{A}$ is the set of links. The sets origin nodes and destination nodes are defined as $\mathcal{R} \subseteq \mathcal{N}$ and $\mathcal{S} \subseteq \mathcal{N}$ respectively. Then let $\mathcal{H} = \mathcal{R} \times \mathcal{S}$ be the set od OD pairs and $\mathbf{q} = \{q_h : h \in \mathcal{H}\}$ the vector of the associated mobility demands. Once again, the demand is considered fix. The set of all path connecting each origine with each destination is represented by $\mathcal{K}$. Regarding network flows, let us define $\mathbf{f} = \{f_k : k \in \mathcal{K}\}$ as the path flow vector and $\mathbf{x} = \{x_a : a \in \mathcal{A}\}$ as the link flow vector. Then based on the network topology, we can establish the link-path incidence matrix $\Delta = \{\delta_{ak} : \delta_{ak} = 1 \text{ if link } a \text{ belongs to path } k \text{ and } \delta_{ak} = 0 \text{ otherwise}\}$. Finally, the OD-path incidence matrix is denoted as $\Xi = \{\xi_{hk} : \text{ if } k \text{ is a path connecting OD pair } h\}$.

The feasible link-path set is then defined as:

$$\Theta = \{(\mathbf{x}, \mathbf{f}) : \mathbf{x} = \Delta\mathbf{f}, \mathbf{q} = \Xi\mathbf{f}, \mathbf{f} \geq 0\} \tag{3.1}$$

based on which the set of feasible link flows can be derived as $\Theta_{\mathcal{A}} = \{\mathbf{x} : (\mathbf{x}, \mathbf{f}) \in \Theta\}$.

Let us associate to the link flow vector a corresponding link travel cost vector $\mathbf{c}(\mathbf{x}) = \{c_a(x_a) : a \in \mathcal{A}\}$. Travel costs $c_a(x_a)$ are considered non negative, continuously differentiable, strictly increasing and fully separable travel cost functions, as formalized in (3.2)-(3.4).

$$c_a(x_a) > 0 \qquad \forall a \in \mathcal{A} \tag{3.2}$$
$$\frac{dc_a(x_a)}{dx_a} > 0 \qquad \forall a \in \mathcal{A} \tag{3.3}$$
$$\frac{dc_a(x_a)}{dx_b} = 0 \qquad a \neq b \tag{3.4}$$

Let $\mathbf{C}(\mathbf{f}) = \{C_k(\mathbf{f}) : k \in \mathcal{K}\}$ be the path travel cost vector. Then it holds that $\mathbf{C} = \Delta' \mathbf{c}$.

The discrete adjustment process based on link flows is expressed as:

$$\mathbf{x}(n+1) = (1 - \lambda(n))\mathbf{x}(n) + \lambda(n)\mathbf{y}(n) \qquad n = 1, 2, \ldots \tag{3.5}$$

where adjustment ratio $\lambda(n) \in (0, 1]$ and $\mathbf{y}(n) = \{y_a(n) : a \in \mathcal{A}\}$ is usually referred to as *target link flow pattern*.

Let us define the set of feasible link flows that decrease the total travel cost with respect to the link costs on day $n$.

$$\Gamma(n) = \Big\{\mathbf{y} : \mathbf{y} \in \Theta_{\mathcal{A}}, \mathbf{y}'\mathbf{c}\big(\mathbf{x}(n)\big) < \mathbf{x}'\mathbf{c}\big(\mathbf{x}(n)\big)\Big\} \tag{3.6}$$

Then if the vector $\mathbf{y}$ satisfies the following condition:

$$\mathbf{y}(n) = \begin{cases} \in \Gamma(n) & \text{if } \Gamma(n) \neq \emptyset \\ = \mathbf{x}(n) & \text{if } \Gamma(n) = \emptyset \end{cases} \tag{3.7}$$

The adjustment process described by (3.5) is referred to as a *discrete rational adjustment process* (DRAP). The DRAP prefigures an adjustment process in which each day a $\lambda(n)$ percentage of users reconsider their travel choices while a $1 - \lambda(n)$ percentage do not reconsider them and adopt the same choices as the day before. This tendency is represented by the target assignment $\mathbf{y}(n)$ that must satisfy condition (3.7). This implies that if there are still available solutions able to decrease the total travel cost, i.e. if $\Gamma(n)$ is a nonempty set, then these will be the solutions adopted. Otherwise, users will reproduce the same choices as the previous day. It should be noted that according to the definition of $\Gamma(n)$ in (3.6), a flow pattern $\mathbf{y}$ is selected if it decreases the network's total travel cost and not necessarily the individual cost. Thus, the process described in (3.5) implies that there may be a portion of users who experience higher travel costs on a given day compared to the previous day. However, the process is described as rational if globally as a result of an adjustment the total travel cost has not increased following an adjustment.

### Path-based formulation

Path-based models represent the most widely adopted class of DTD assignments for two main reasons: (1) they allow a more flexible definition of flow adjustment processes that occur on a day-to-day basis; and (2) user choice behavior models are more easily represented when the set of choices is defined on paths rather than links. This is because usually an individual evaluates a link not singularly but in relation to the path of which it is a portion.

Regarding DTD path-based models, the work by Smith (1984) is of substantial relevance. The author proposed a *proportional switch adjustment process* that take place whenever the system is at a point of disequilibrium and due to which it tends to return to an equilibrium state. The process is called proportional switch because each day the flows switch from one path to another proportionally to the travel cost difference between them. The flow previously on the expensive path is redistributed proportionally favoring those paths that provide greater travel cost savings. The original model has been adopted in traffic assignment problems (Huang and Lam, 2002; Mounce and Carey, 2011), in the definition of traffic control policies (Guo, 2013; Liu and Smith, 2015) and has also been modified to accommodate various user choice behavior rules (Guo, 2013; Smith and Watling, 2016).

In this section, the general framework for path-based DTD assignment models proposed by Guo and Huang (2016) is presented. The notation is the same as that used in the previous section but with the addition of two additional *decision-making travel cost vectors* associated to links and paths. Similarly to the approach followed in section section 2.5, let us define $f_k^h = f_k : \xi_{hk} = 1$ and $C_k^h = C_k : \xi_{hk} = 1$ where $\xi_{hk}$ is an element of the OD-path incidence matrix.

Let $\bar{\mathbf{c}}(\mathbf{x}) = \{\bar{c}_a(\mathbf{x}) : a \in \mathcal{A}\}$ be the decision-making link travel cost vector and $\bar{\mathbf{C}}(\mathbf{f}) = \{\bar{C}_k^h(\mathbf{f}) : k \in \mathcal{K}, h \in \mathcal{H}\}$ as the decision-making path cost vector. Then, $\bar{\mathbf{C}}(\mathbf{f}) = \Delta' \bar{\mathbf{c}}(\mathbf{x})$ holds. Decision making travel costs are the costs according to which users make their decisions and while they are still flow-dependent they may be different from the actual travel costs. Furthermore, the following conditions hold: $c_a(\mathbf{x}) > 0$, $C_k^h(\mathbf{f}) > 0$, $\bar{c}_a(\mathbf{x}) > 0$ and $\bar{C}_k^h(\mathbf{f}) > 0$. Non-separable link cost functions are considered, i.e. the travel cost associated with a link may also depend on the flows on other links.

Using the above notation, the general path-based dynamical system is defined as:

$$\mathbf{f}(n+1) = \mathbf{f}(n) + \tau(n)\boldsymbol{\Lambda}(\mathbf{f}(\mathbf{n})) \qquad n = 1, 2, \ldots \quad (3.8)$$

where $n$ refers to the $n$-th day, the adjustment parameter $\tau(n) > 0$ and $\boldsymbol{\Lambda}(\mathbf{f}) = \{\Lambda_{kh}(\mathbf{f}(n)) : k \in \mathcal{K}, h \in \mathcal{H}\}$. All the elements of $\boldsymbol{\Lambda}(\mathbf{f}(\mathbf{n}))$ are governed by:

$$\Lambda_{kh}(\mathbf{f}(n)) = \sum_{s \in \mathcal{K}_h} \left[\alpha_{sh}(n)\gamma_{sk}^h(\mathbf{f}(n)) - \alpha_{kh}(n)\gamma_{ks}^h(\mathbf{f}(n))\right] \quad k \in \mathcal{K}, h \in \mathcal{H} \quad (3.9)$$

where $\mathcal{K}_h \subseteq \mathcal{K}$ is the set of path connecting OD pair $h$. Function $\gamma_{ks}^h(\mathbf{f})$ is defined as:

$$\gamma_{ks}^h(\mathbf{f}) = \max\{\bar{C}_k^h(\mathbf{f}) - \bar{C}_s^h(\mathbf{f}), 0\} \qquad k \in \mathcal{K}, s \in \mathcal{K}, h \in \mathcal{K} \quad (3.10)$$

while $\boldsymbol{\alpha}(n) = \{\alpha_{kh}(n) : k \in \mathcal{K}, h \in \mathcal{H}\}$ is the vector of flow transfer parameters. Each element $\alpha_{kh}(n)$ scales the flows that on day $n+1$ switch from path $k$ to other paths of the same OD pair $h$ and they satisfy the following conditions.

$$\alpha_{kh}(n) \geq 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (3.11)$$

$$\tau(n)\alpha_{kh}(n) \sum_{s \in \mathcal{K}_h} \gamma^h_{ks}(\mathbf{f}(n)) \leq f^h_k(n) \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (3.12)$$

$$\alpha_{kh}(n) = 0 \Leftrightarrow f^h_k(n) = 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (3.13)$$

For each day $n$, condition (3.11) is a non-negativity constraint. Condition (3.12) implies that the total amount of flow leaving path $k$ is not larger than the flow actually on the path thus ensuring that path flows cannot be negative after an adjustment. Finally, condition (3.13) implies that the flow transfer parameter associated with path $k$ of the OD $h$ is positive if and only if the corresponding path flow $f^h_k(n)$ is positive.

Guo and Huang (2016) demonstrate that the process described in (3.8) satisfies flow conservation and that, depending on the form assumed by the decision-making cost functions $\bar{\mathbf{C}}(\mathbf{f})$, it converges to a set of equilibrium such as the Wardrop UE (Wardrop, 1952), the Logit-based SUE (Daganzo and Sheffi, 1977) or the Bounded Rationality User Equilibrium (BRUE). The concept of BRUE is detailed in section 4.1.

# Chapter 4

# Bounded Rationality Models

The perfect rationality of the decision maker is a widely leveraged assumption in decision-making models and the transportation domain is no exception. Typically, individuals are represented as utility maximizers (or disutility minimizers). In contrast to this view, an alternative has been proposed by Simon (1957) suggesting that the individual is generally rationally bounded while making decisions and, rather than aiming at the optimal solution, he/she most often seeks to avoid the worst.

The reasons behind this are multiple and range from the inability to get access to all the information needed to formulate a decision to the fact that it is difficult for the individual, even with all the necessary information, to actually choose the best solution because of the complexity of the circumstances in which he or she has to make the choice. Furthermore, the search for a solution is a dynamic process. The alternatives are discarded until a satisfactory one is found, but in many cases the individual may not know whether a better one may still exist. Finally, as shown by Simon (1957), the longer the search process lasts, the more the individual's aspirational level regarding the outcome of his or her decision tends to adjust so that an acceptable solution can be found.

In the transportation field, the concept of bounded rationality (BR) was introduced in an attempt to explain several user travel choice behaviors observed by experimental analyses that were in contrast to the more classical perfect rationality (PR) models. For the first time in the transportation science, Mahmassani and Chang (1987) adopted the concept of BR to model users pre-trip departure time choice behavior in a network with a single bottleneck. Since this early work, a growing body of literature has been developed where the BR concept is incorporated into traffic assignment models (Fonzone and Bell, 2010), DTA models (Szeto and Lo, 2006), traffic policy making (Marsden et al., 2012), transportation planning (Gifford and Checherita, 2007; Khisty and Arslan, 2005) and transport safety (Sivak, 2002).

The reasons behind why users do not always choose the shortest or least expensive route can be attributed to multiple factors. Some of these relate to road characteristics such as distance, number of intersections, complexity or aesthetics of the route, as well as the presence or absence of freeways. These are all factors that might motivate a traveler to choose one alternative over another (Ramming, 2001; Bekhor, Ben-Akiva, and Ramming, 2006; Papinski, Scott, and Doherty, 2009; Bovy and Stern, 2012; Prato, Bekhor, and Pronello, 2012). Other similarly important factors may be the completeness and reliability of the information available to users (Ben-Elia and Shiftan, 2010; Bifulco, Di Pace, and Viti, 2014; Ma, Wu, and Wang, 2014; Zhang et al., 2015b; Chunyan et al., 2016), socio-economic attributes (Tawfik and Rakha, 2013) and road safety (Kusakabe, Sharyo, and Asakura, 2012). On the other hand, suboptimal alternative choice selection can also be explained on the basis of perceptive limitations and cognitive biases intrinsic within the individual's choice process. In their comprehensive review, Di and Liu (2016) collected and discussed several major empirical evidence in literature showing how the assumption of perfect rationality may result in an idealized representation of a traffic system. They are briefly summarized here.

**Heuristic, bias and cognitive limit:** Being in the same situation, the individual may react differently depending on how the problem is "framed" (Tversky and Kahneman, 1985). Biases related to the choice process generally do not vanish even after a learning period and even if incentive or punishment mechanisms are provided demonstrating how the biases are "substantial" and recur with regularity (Conlisk, 1996). In addition, individuals adopt heuristic choice processes because the tradeoff between the goodness of a solution and the efficiency of the decision process are both evaluated. An optimal choice making process may require a lot of time and effort in both information retrieval and choice selection. Hiraoka et al. (2002) showed that drivers prefer those routes that provide less travel time, require less cognitive effort and provide lower stress levels. However, among this three factors the amount of cognitive effort required for route planning is the one that most influences the decision.

**Nonexistence of perfect rationality via learning process:** It is well known within the economic literature that a learning process can improve the individuals' rationality. However, Conlisk (1996) showed that learning can not only improve individual's decision-making process but can also induce habit phenomena when past winning choices are more likely to be replicated in the present. Within route choice models, Lotan (1997) pointed out how drivers who are familiar with the network tend to select the same roads even when they are no longer the shortest ones. In contrast, unfamiliar drivers tend to choose roads more dynamically and rely more on real-time traffic information (Hiraoka et al., 2002).

**Habit and inertia:** Relying on past experiences is a common practice exploited by people within choice-making mechanisms because it is a strategy that has the potential to save a significant amount of cognitive energy (Samuelson and Zeckhauser, 1988). However, when circumstances are not sufficiently stable, this strategy degenerates into an irrational reliance on habit that prevents the search for new solutions. Cantillo, Ortuzar, and Williams (2007) showed how not considering the influence of inertia in a modal choice model can result in overestimating the benefits associated with an infrastructure investment. Also, Carrion and Levinson (2012) research on commuters' habits leads to this conclusion. The authors showed that commuters tend to change their habitual route exclusively when travel cost exceeds a certain threshold. Such evidence has been formalized in the concept of *indifference band* in a series of fundamental works by Mahmassani and his colleagues (Mahmassani and Liu, 1999; Srinivasan and Mahmassani, 1999).

**Myopia:** Individuals tend not to consider the long-term consequences arising from choices they are making in the present. Analogously regarding travel choices, users tend to switch towards routes based on a promising immediate time savings which may result in an overall more expensive trip. Bogers, Viti, and Hoogendoorn (2005) showed that closest past experiences are most likely to influence present choices.

This is due to individuals' limited memory capacities, which in part involve users' myopia.

More recently, several empirical studies have confirmed how travelers sometimes do not choose the shortest route in order to reach their destination. Zhu and Levinson (2015) collected GPS data from 143 residents over a 13-week period. The results show that only 40% of the trips take place on minimum time routes while 90% of the times the selected routes are no more than 5 minutes longer than che shortest one. Similarly, exploiting the GPS trajectories of 20000 taxis in Shenzhen (China), Yildirimoglu and Kahraman (2018) evaluate the divergence between objective measures and classical equilibrium assumptions finding that most of the trips (62%) are not at minimum travel time. Similar results have been observed in the studies by Vreeswijk et al. (2013) and Hadjidimitriou et al. (2015). In the first study, 20 participants' travel habits were collected over a period of 20 days from an experiment in Virginia (United States). The authors found that on average only in 74% of the cases the shortest route was selected, although the data varies considerably among OD pairs. In the second study, exploiting the GPS data from more than 14000 recurring trips in the Italian province of Reggio Emilia (Italy), the results show that around 25% of users choose the shortest path and as a consequence travelers spent 30% more time than necessary on the network. In addition, it is shown that users are willing to accept even significant increases in travel time as long as the route is reliable (small travel time fluctuations). Finally, González Ramírez (2020) conducted a simulation game experiment where 496 participants were asked to select routes connecting 41 OD pairs presented on a road map of the city of Lyon (France). In 71% of the cases, users were provided with an estimation of travel times in an attempt to avoid travel time perception bias. The collected dataset included 5535 route choices. It was found that despite information provision only 60.5% of travel choices resulted in the selection of a shortest route, suggesting how in a real-world setting, where information may not be complete, this percentage may rise further. Even more interestingly, the authors find that users evaluate differences in travel time relatively rather than in absolute terms. In other words, what matters is the increase in travel time compared to an average level. The longer the routes, the more tolerant users are to even significant travel time variations as they affect the overall trip travel time by a relatively small percentage. Furthermore, the results show that user indifference appears to be heterogeneous with an average indifference band of about 31% of the travel time of the shortest route. In order to estimate indifference bands, the author proposes a bounded rational choice set generation mixed logit model (BCRS) characterized by a two-step process which firstly allows the preselection of the set of attractive paths having travel times within a certain indifference band, extracted by a Weibull distribution. Among these, the chosen route is then selected employing a rational route choice process.

## 4.1 Bounded Rationality DTD models

The phenomenon of bounded rationality has been repeatedly detected in the field of transportation where it manifests in the form of a shortest path violation. In several studies it is estimated that only between 60% to 90% of drivers actually choose the shortest path (Bekhor, Ben-Akiva, and Ramming, 2006; Prato and Bekhor, 2006; Zhu, 2010). The analysis of the residual gap in flow patterns between the pre-collapse situation of the I-35W Mississippi River bridge and the post-reconstruction scenario has been a fundamental case study in this line of research (Zhu, 2010), showing that significant infrastructure changes can lead to permanent flow shifts. Furthermore, the series of studies conducted by Srinivasan and Mahmassani (1999) showed how the shortest route choice violation persists even in case drivers are provided with all the information they need to formulate the decision.

The definition of Bounded Rationality User Equilibrium (BRUE) (Mahmassani and Chang, 1987), the state at which each user cannot significantly decrease (i.e. beyond a certain threshold) his travel cost by unilaterally deviating, has been extensively addressed within the framework of static traffic assignments (Lou, Yin, and Lawphongpanich, 2010; Di et al., 2014). In recent years, the concept of bounded rationality has also spread to the research field of DTD models. Following the study on the shift in traffic patterns due to the collapse of the I-35W Mississippi River bridge, some research works have pointed out that these phenomena are poorly represented by traditional day-to-day models (He and Liu, 2012; Guo and Liu, 2011a; Di et al., 2015). More specifically, Guo and Liu (2011a), mutating the idea behind the work of He, Guo, and Liu (2010), a day-to-day link-based assignment model is proposed which embeds the features of bounded rationality and it is shown that, although the dynamics converges to a set of BRUEs, the system does not necessarily tend to the original one after a perturbation. In other words, due to bounded rationality, the equilibrium state of the network can change irreversibly following a disturbance. The same phenomenon is observed by Di et al. (2015), where stability properties for bounded rationality dynamics are provided. Wu et al. (2013) incorporated the features of bounded rationality into a cost updating mechanism originally proposed by Cantarella and Cascetta (1995) and the dynamics are applied to the railway context. Ye and Yang (2017) adopted the bounded rationality concepts expressed through indifference band representation within day-to-day proportional-switch adjustment and tatonnement dynamics. Shang et al. (2017) investigate the impact of information sharing among travelers through the use of an agent-based day-to-day assignment model. The user's adjustment process is therefore influenced also by the experience gathered by other users within the same "information-cluster", whose generation is obtained by employing percolation theory. The process convergence to a steady state is then investigated under both scenarios with or without user bounded rationality. Generalizing the original concept of proportional switch process by Smith (1984), Guo and Huang (2016) formulated a general framework for day-to-day path-based processes which can accomodate a variety of route choice behaviors. Depending on certain preconditions, the model evolves towards the classical UE and SUE or towards a BRUE. Furthermore, Zhang et al. (2018) suggested a non-linear pairwise adjustment process in both absolute and relative bounded rationality implementations. In the first case, the indifference band is constant and unique for each OD pair, while in the second case it scales with the actual travel costs and thus indirectly with the level of congestion and size of the network. Finally, Zhang et al.

(2019) extended the model by incorporating net marginal cost gain in the route choice process. This means that users, when considering a switch, not only take into account the current absolute cost differences between the route they are on and the one they will potentially switch to, but they also rely on the cost increase resulting from their choice influenced by path marginal costs. In other words, both the cost difference between the paths moderated by a variable indifference band, which scales with respect to the costs, and the net marginal cost, intended as the marginal cost difference between the paths, are taken into account in the user decision process. According to Zhang et al. (2019), the users prefer, given the same amount of actual travel cost savings, the routes that will result in a lower cost increase once the switch takes place. As a result, given the same travel cost savings, users tend to prefer those routes that overlap with the one currently used, since the overlapped portion does not contribute to the net marginal cost increase.

## 4.2 Boundedly Rational User Equilibrium

The concept of Boundedly Rational User Equilibrium (BRUE) was first proposed by Mahmassani and Chang (1987) through the concept of indifference band applied to the departure-time choice. The underlying assumption is that users are not sensitive to minute changes in travel costs. For this reason, the solutions adopted are not necessarily those with the lowest possible cost but may likewise have a cost that diverges from the minimum by no more than a certain value. The user then changes choice exclusively when the stimulus intensity exceeds a certain threshold represented by an indifference band. A rigorous mathematical formulation of the BRUE for route choice behavior was proposed by Lou, Yin, and Lawphongpanich (2010).

Alternatively to this definition, Zhang and Yang (2015) proposed the concept of *Inertial User Equilibrium* (IUE). The concept of bounded rationality in this case is represented by a set of prevailing paths, i.e. the inertial user equilibrium. More specifically, different groups of users can only use a subset of all available paths for their OD pair which is defined as the inertial set. Therefore, a flow is a IUE if within each inertial set Wardrop's first principle is satisfied.

### Boundedly Rational User Equilibrium

Let represent a traffic network by means of an oriented graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is the set on nodes and $\mathcal{A}$ is the set of paths. Two subset of nodes are defined as origin nodes $\mathcal{R} \subseteq \mathcal{N}$ and destination nodes $\mathcal{S} \subseteq \mathcal{N}$. Let $\mathcal{H} = \mathcal{R} \times \mathcal{S}$ be the set of OD pairs associated with the mobility demand vector $\mathbf{q} = \{q_h : h \in \mathcal{H}\}$. The set of all loop-free path connecting an origin node with a destination node is denoted by $\mathcal{K}$. It is then possible to establish a relation beetween path and links by means of the link-path incidence matrix $\Delta = \{\delta_{ak} : a \in \mathcal{A}, k \in \mathcal{K}\}$ where $\delta_{ak} = 1$ if link $a$ belongs to path $k$. In order to associate paths with their respective OD pairs, let us define the OD-path incidence matrix $\Xi = \{\xi_{hk} : h \in \mathcal{H}, k \in \mathcal{K}\}$ where each element $\xi_{hk} = 1$ if path $k$ connect the origin node to the destination node of the OD pair $h$. Regarding traffic flows, let $\mathbf{x} = \{x_a : a \in \mathcal{A}\}$ and $\mathbf{f} = \{f_k : k \in \mathcal{K}\}$ be the link flow vector and the path flow vector respectively. The corresponding travel cost vectors are denoted by $\mathbf{c} = \{c_a : a \in \mathcal{A}\}$ for the links and $\mathbf{C} = \{C_k : k \in \mathcal{K}\}$ for the paths.

Similarly to what done in section 2.5, let us state that a pattern of flows $(\mathbf{x}, \mathbf{f})$ is feasible if it falls within the set $\Theta$ as defined in (2.11).

Finally, let us defined two vectors associated with the OD pairs. Let $\boldsymbol{\pi} = \{\pi_h : h \in \mathcal{H}\}$ be the minimum cost vector where each element $\pi_h$ is the absolute minimum path cost for OD pair $h$ and $\boldsymbol{\epsilon} = \{\epsilon_h : h \in \mathcal{H}\}$ be the threshold vector where each element $\epsilon_h$ represent the threshold value for all users of OD pair $h$.

It is therefore possible to provide the following two definitions.

**Definition 3.** *A path is "acceptable" if the difference between its travel time or cost and that of the shortest or least-cost path is no larger than a pre-specified threshold value.*

**Definition 4.** *A path flow distribution is a BRUE if it is feasible and compatible with the travel demands (i.e. falls within $\Theta$) and every user uses an acceptable path.*

The mathematical formulation of a BRUE flow pattern is defined as follows.

$$C_k \geq \Xi'_k \boldsymbol{\pi} \qquad \forall k \in \mathcal{K} \tag{4.1}$$

$$f_k > 0 \implies C_k \leq \Xi'_k (\boldsymbol{\pi} + \boldsymbol{\epsilon}) \qquad \forall k \in \mathcal{K} \tag{4.2}$$

where $\Xi'_k$ is the $k$-th row of the transposed OD-path incidence matrix, therefore $\Xi'_k \boldsymbol{\pi} = \pi_h$ and $\Xi'_k \boldsymbol{\epsilon} = \epsilon_h$ hold if $k$ is a path connecting OD pair $h$. Let us define $f_k^h = f_k : \xi_{hk} = 1$ and $C_k^h = C_k : \xi_{hk} = 1$, then (4.1)-(4.2) can be reformulated as follows:

$$C_k^h \geq \pi_h \qquad \forall k \in \mathcal{K}, h \in \mathcal{H} \tag{4.3}$$

$$f_k^h > 0 \implies C_k^h \leq \pi_h + \epsilon_h \qquad \forall k \in \mathcal{K}, h \in \mathcal{H} \tag{4.4}$$

where (4.3) implies that all path cost for an OD pair $h$ must be greater than the minimum one $\pi_h$ while (4.4) establish that each used path must have a travel cost that is larger than the minimum by no more than $\epsilon_h$.

By introducing a vector of slack variables $\tilde{\boldsymbol{\epsilon}} = \{\tilde{\epsilon}_h : h \in \mathcal{H}\}$, it is possible to summarize the above conditions similarly to (2.17)-(2.19) and (2.20)-(2.22).

$$\mathbf{C} - \Xi'(\boldsymbol{\pi} + \tilde{\boldsymbol{\epsilon}}) = 0 \tag{4.5}$$

$$\mathbf{f} \circ \Xi'(\boldsymbol{\epsilon} - \tilde{\boldsymbol{\epsilon}}) \geq 0 \tag{4.6}$$

$$\tilde{\boldsymbol{\epsilon}} \geq 0 \tag{4.7}$$

$$(\mathbf{x}, \mathbf{f}) \in \Theta \tag{4.8}$$

where operator ∘ is an Hadamard (element-wise) product. The same conditions can be expressed alternatevely using the notation in (4.3)-(4.4).

$$C_k^h - \pi_h - \tilde{\epsilon}_h = 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (4.9)$$

$$f_k^h(\epsilon_h - \tilde{\epsilon}_h) \geq 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (4.10)$$

$$\tilde{\epsilon}_h \geq 0 \qquad h \in \mathcal{H} \quad (4.11)$$

$$(\mathbf{x}, \mathbf{f}) \in \Theta \quad (4.12)$$

Note that if the vector of thresholds is equal to zero, i.e. when $\pi_h = 0$ for each OD pair, conditions (4.9)-(4.12) reduce to the (2.20)-(2.22) and the BRUE collapses into a classical UE. Assuming $\pi_h = 0 \; \forall h \in \mathcal{H}$ condition (4.10) reduces to:

$$f_k^h \tilde{\epsilon}_h \geq 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (4.13)$$

and once noted from the (4.9) that $\tilde{\epsilon}_h = C_k^h - \pi_h$, conditions (4.10) and (4.11) reduce to

$$f_k^h(C_k^h - \pi_h) \geq 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (4.14)$$

$$C_k^h - \pi_h \geq 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (4.15)$$

which are the UE conditions. A BRUE is not necessarily unique, and for this reason we refer to a BRUE set of flow patterns that satisfy conditions (4.9)-(4.12). It should also be noted that a UE flow pattern falls necessarily within a BRUE set. In fact, if a flow pattern satisfies conditions (4.14)-(4.14) then it necessarily also satisfies (4.9)-(4.11).

### Inertial User Equilibrium

The concept behind the Inertial User Equilibrium (IUE) proposed by Zhang and Yang (2015) is that users use only a subset of all available paths for their pair. Such a set of paths is called an inertial set. Within each OD pair, there can be different groups of users each using their own inertial set of paths among those potentially available. Let us apply the same notation as in the previous section with the following additions necessary to describe the inertial sets.

Let $\mathcal{K}_h \subseteq \mathcal{K}$ be the set of all available paths for OD pair $h \in \mathcal{H}$. Let $\mathbf{W}_h = \{W \subseteq \mathcal{K}_h : h \in \mathcal{H}\}$ be the set of all inertial patterns for OD pair $h$. Each element within $\mathbf{W}_h$ is a particular sub-set of all available paths for the OD pair h. A probability vector is then defined for each OD pair denoted by $\mathbf{p}_h = \{p_h^W : h \in \mathcal{H}, W \in \mathbf{W}_h\}$ where an element $p_h^W$ represents the probability that a traveler of OD pair $h$ will use a path within the inertial set $W \in \mathbf{W}_h$. The following conditions must be satisfy.

$$p_h^W \geq 0 \qquad W \in \mathbf{W}_h, h \in \mathcal{H} \quad (4.16)$$

$$\sum_{W \in \mathbf{W}_h} p_h^W = 1 \qquad h \in \mathcal{H} \quad (4.17)$$

The following two definitions are provided.

**Definition 5.** *A traveler with demand between OD pair $h \in \mathcal{H}$ has inertia pattern $W$ when $W$ is a subset of $\mathcal{K}_h$, i.e. $W \subseteq \mathcal{K}_h$ , and the traveler only has his or her route choices within the choice set $W$.*

**Definition 6.** *Consider a traveler with demand between OD pair $h \in \mathcal{H}$ and inertia pattern $W$. The traveler is fully inertial if $|W| = 1$; the traveler is non-inertial or fully non-inertial if $W = \mathcal{K}_h$; the traveler is partially inertial otherwise. A traveler is inertial if he or she is fully inertial or partially inertial.*

Definition 5 formalizes the concept of inertial pattern while definition 6 differentiates among different groups of users. When $|W| = 1$, a traveller will chose the same path independently from the congestion level. At the opposite, $W = \mathcal{K}_h$ indentify a group of users who will eventually exploit all the available path for OD pair $h$ if required.

Regarding path flows, let $f_k^{hW}$ be the portion of flow on path $k \in \mathcal{K}_h$ generated by inertia pattern $W \in \mathbf{W}_h$. Then we have

$$f_k^h = \sum_{W \in \mathbf{W}_h} f_k^{hW} \qquad k \in \mathcal{K}_h, h \in \mathcal{H} \quad (4.18)$$

$$q_h^W = q_h p_h^W \qquad W \in \mathbf{W}_h, h \in \mathcal{H} \quad (4.19)$$

$$q_h^W = \sum_{k \in \mathcal{K}_h} f_k^{hW} \qquad W \in \mathbf{W}_h, h \in \mathcal{H} \quad (4.20)$$

where (4.18) establish that the flow associated to a path $k$ of an OD pair $h$ is determined by the contribution of the flows generated by each inertial pattern $W \in \mathbf{W}_h$ while (4.19) identifies the mobility demand associated with each inertial pattern and (4.20) is a flow conservation equation.

It is then possible to provide a formal definition of inertial user equilibrium.

**Definition 7.** *The traffic flow is at inertial user equilibrium when it follows Wardrop's first principle in every inertia pattern $W \in \mathbf{W}_h$ for every OD pair $h \in \mathcal{H}$.*

Let $\boldsymbol{\pi} = \{\pi_h^W : W \in \mathbf{W}_h, h \in \mathcal{H}\}$ be a vector of minimum costs where each element $\pi_h^W$ is the minimum cost associated to the inertia patter $W \in \mathbf{W}_h$. The mathematical formulation of a IUE flow pattern is defined as follows.

$$f_k^{hW}(C_k^h - \pi_h^W) \geq 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (4.21)$$

$$C_k^h - \pi_h^W \geq 0 \qquad k \in \mathcal{K}, h \in \mathcal{H} \quad (4.22)$$

$$(\mathbf{x}, \mathbf{f}) \in \Theta \quad (4.23)$$

Conditions (4.21)-(4.23) are equivalent to those for the UE but detailed for each inertial pattern. Note that, when referring to path costs $C_k^h$, there is no superscript $W$ because, regardless of the inertial pattern, the cost on a path $k$ is fixed given a particular configuration of flows on the network.

# Part II

# Dynamic traffic assignment models for disrupted networks

## Chapter 5

# A progressive traffic assignment model for disrupted network

Based on *Siri, Siri, and Sacone (2020b)*

In this chapter, a day-to-day link-based traffic assignment model design in order to represent the evolution of a traffic network when subject to a disruption is presented. The kind of disruption considered here is a sudden and significant alteration of the network infrastructure topology. The model then estimates the progressive readjustment of traffic flows onto the network until a new equilibrium is reached.

The assignment procedure is characterized by a series of progressive UE traffic assignments performed under variable conditions taking into account the habits of users and the inertia of the system as a whole. The main ideas behind the model are two: the former concerns the specific behavior of users, the latter regards the dynamics of the system as a whole. Regarding the behavior of users, the model accounts for the fact that they consider roads that are alternative to those already in use only when the travel costs they experience have increased significantly. By "significantly" we mean more than a certain percentage value expressed in the model by the *user tolerance index* $\Omega$, which will be defined rigorously in the following. For each iteration, the proposed algorithm, comparing for each origin-destination pair the increase in travel cost that the users experience with the maximum threshold they tolerate, verifies whether they are satisfied or not with the current situation. If they are not, in the next iteration they will be assigned considering a slightly larger set of paths. The assignment that is obtained from this process is defined as *target assignment* and represents at each iteration the direction towards which the system tends to move. Therefore, considering the dynamics of the system as a whole, it is not expected to jump instantly from one new solution to another, but because of its intrinsic inertia, represented by the *inertia coefficient* $\beta$, the system should evolve through a series of states that are somewhere in between the best possible assignment given the current circumstances and the previous one.

The model proposed in this chapter can be considered a dynamic variant of the static IUE assignment model proposed by Zhang and Yang (2015) (discussed in section 4.2). However, in this case each OD pair is associated with only one inertial set, i.e. all users of the same OD pair can use only one specific set of paths that does not necessarily include all those available to the pair. Moreover, such path sets are not fixed but they may vary over time, meaning new paths are considered and added to the sets if users are not satisfied with current travel costs.

This chapter is organized as follows. In section section 5.1 the notation is presented. In section section 5.2 the model is described in details. Further in section section 5.3 some performance indices are defined while in section section 5.4 a viable implementation of the model is proposed. Finally in section section 5.5 some conclusions are drawn.

## 5.1 Notation

The notation used in the model is as follows. First of all, the topological quantities and sets are defined:

- $G(\mathcal{N}, \mathcal{A})$: graph denoting the transportation network consisting of a set of nodes $\mathcal{N}$ and directed arcs $\mathcal{A}$, where $|\mathcal{A}| = A$
- $\mathcal{R} \subseteq \mathcal{N}$ and $\mathcal{S} \subseteq \mathcal{N}$: set of origin and destination nodes respectively
- $\mathcal{H} = \{h : h \in \mathcal{R} \times \mathcal{S}\}$: set of all origin-destination pairs
- $\mathcal{K}_h^n$=set of all available paths for OD pair $h \in \mathcal{H}$ on day $n$
- $\mathcal{L}_h^{\text{UE}}$: set of paths, obtained through the UE traffic assignment, that are actually used by each origin-destination pair $h \in \mathcal{H}$ at the equilibrium before the occurrence of the disruption
- $\mathcal{L}_h^{\text{SA}}$: set of path actually used by users of OD pair $h \in \mathcal{H}$ just after the disruption
- $\mathcal{L}_h^n$: set of path for OD pair $h \in \mathcal{H}$ used for the assignment on each day $n$

Traffic flows and traffic assignment variables are defined as follows:

- $\mathbf{q} = \{q_h : h \in \mathcal{H}\}$: traffic demand vector, where $q_h$ is the traffic demand associated with OD pair $h$
- $\mathbf{x}(n) = \{x_a(n) : a \in \mathcal{A}\}$: link-flow vector on day $n$, where $x_a(n)$ is the flow on link $a$ on day $n$

- $\mathbf{x}^{\text{UE}} = \{x_a^{\text{UE}} : a \in \mathcal{A}\}$: link-flow vector at the equilibrium before the disruption
- $\mathbf{x}^{\text{SA}} = \{x_a^{\text{SA}} : a \in \mathcal{A}\}$: Shock Assignment link-flow vector immediately after the disruption
- $\mathbf{z}(n) = \{z_a(n) : a \in \mathcal{A}\}$: target link-flow vector computed for day $n$
- $\mathbf{x}^{\text{FE}} = \{x_a^{\text{FE}} : a \in \mathcal{A}\}$: Final link-flow vector representing the new equilibrium reached by the system after the disruption
- $\mathbf{f}(n) = \{f_k^h(n) : k \in \mathcal{L}_h^n, h \in \mathcal{H}\}$: path-flow vector on day $n$, where $f_k^h(n)$ is the traffic flow on path $k \in \mathcal{L}_h^n$ of OD pair $h \in \mathcal{H}$ on day $n$

The travel costs on links and paths are defined as follows:

- $\mathbf{c}(\mathbf{x}(n)) = \{c_a(x_a(n)) : a \in \mathcal{A}\}$ Link travel cost vector on day $n$, where $c_a(x_a(n))$ is a non negative, monotonically increasing, fully separable travel cost function of link $a \in \mathcal{A}$ on day $n$
- $\mathbf{C}(\mathbf{f}(n)) = \{C_k^h(\mathbf{f}(n)) : k \in \mathcal{L}_h^n, h \in \mathcal{H}\}$: path travel cost vector on day $n$, where $C_k^h(\mathbf{f}(n))$ is the travel cost on path $k \in \mathcal{L}_h^n$ of OD pair $h \in \mathcal{H}$ on day $n$
- $TTC(n)$: network total travel cost on day $n$

Finally, the coefficients of the model are:

- $\Omega \in [0, +\infty)$: *user tolerance index*
- $\beta \in [0, 1]$: *inertia coefficient*

The notation introduced in this section is summarized in Table 5.1.

## 5.2 The Progressive Traffic Assignment Model

In the following, the main elements of the assignment model are illustrated, according to the flow chart of fig. 5.1.

| Sets | |
|---|---|
| $\mathcal{N}$ | Set of nodes |
| $\mathcal{A}$ | Set of links |
| $\mathcal{R}$ | Set of origin nodes |
| $\mathcal{S}$ | Set of destination nodes |
| $\mathcal{H}$ | Set of origin-destination (OD) pairs |
| $h$ | OD couple $h \in \mathcal{H}$ |
| $\mathcal{K}_h^n$ | Set of all loop-free paths for OD pair $h \in \mathcal{H}$ on day $n$ |
| $\mathcal{L}_h^{\mathrm{UE}}$ | Set of paths actually used by each origin-destination pair $h \in \mathcal{H}$ at the equilibrium before the disruption |
| $\mathcal{L}_h^{\mathrm{SA}}$ | Set of path actually used by users of OD pair $h \in \mathcal{H}$ just after the disruption |
| $\mathcal{L}_h^n$ | Set of path for OD pair $h \in \mathcal{H}$ used for the assignment on each day $n$ |
| **Flow Vectors** | |
| $\mathbf{q}$ | Travel demand vector |
| $q_h$ | Travel demand of OD pair $h \in \mathcal{H}$ |
| $\mathbf{x}(n)$ | Link-flow vector on day $n$ |
| $x_a(t)$ | Flow on link $a \in \mathcal{A}$ on day $n$ |
| $\mathbf{x}^{\mathrm{UE}}$ | Link-flow vector at the equilibrium before the disruption |
| $\mathbf{x}^{\mathrm{SA}}$ | Link-flow vector immediately after the disruption |
| $\mathbf{z}(n)$ | Link-flow target assignment for day $n$ |
| $z_a(n)$ | Target flow on link $a \in \mathcal{A}$ on day $n$ |
| $\mathbf{x}^{\mathrm{FE}}$ | Final link-flow vector at the new equilibrium |
| $\mathbf{f}(n)$ | Path-flow vector on day $n$ |
| $f_k^h(t)$ | Flow on path $k \in \mathcal{L}_h^n, h \in \mathcal{H}$ on day $n$ |
| **Travel Costs** | |
| $\mathbf{c}\big(\mathbf{x}(n)\big)$ | Link travel cost vector on day $n$ |
| $c_a\big(x_a(n)\big)$ | Travel cost on path $a \in \mathcal{A}$ on day $n$ |
| $\mathbf{C}\big(\mathbf{f}(n)\big)$ | Path travel cost vector on day $n$ |
| $C_k^h\big(\mathbf{f}(n)\big)$ | Travel cost on path $k \in \mathcal{L}_h^n, h \in \mathcal{H}$ |
| $TTC(n)$ | Newtork total travel cost on day $n$ |
| **Parameters** | |
| $\Omega$ | User tolerance index |
| $\beta$ | Inertia coefficient |

TABLE 5.1: Main notation

1) *User-Equilibrium Assignment*. First of all, given a the mobility demand $\mathbf{q}$ which is fixed and link cost functions $c_a(x_a)$, which are nonnegative, monotonically increasing, and fully separable, an UE assignment is performed as described in Problem 1. This allows to obtain the pattern of traffic flows $\mathbf{x}^{\text{UE}}$ at the equilibrium and the associated link travel costs vector $\mathbf{c}(\mathbf{x}^{\text{UE}})$. Then let the set of paths potentially used by users of each OD pair at the equilibrium before the disruption be defined as follows

$$\mathcal{L}^{\text{UE}} = \{k \in \mathcal{K}_h : C_k^h = \pi^h\} \tag{5.1}$$

where $\pi^h$ is the minimum travel cost for OD pair $h \in \mathcal{H}$. Equation (5.1) establish that a path in order to be used by users needs to have minimum travel cost. Finally, be $\mathbf{C}(\mathbf{f}^{\text{UE}})$ the associated path travel cost vector where $\mathbf{f}^{\text{UE}}$ is a path flow vector consistent with the link flow vector $\mathbf{x}^{\text{UE}}$ computed by the UE traffic assignment. For the two vectors to be consistent, the following relationship must hold

$$\mathbf{x}^{\text{UE}} = \Delta \mathbf{f}^{\text{UE}} \tag{5.2}$$

where $\Delta$ is the link-path incidence matrix.

2) *Shock-Assignment*. Once the disruption has occurred, a so-called Shock-Assignment is performed. Only the flows on the routes directly involved are reassigned by means of all-or-nothing assignment. At this moment only those users whose path has been directly affected by the disruption reorganize themselves, having no alternatives. All the others suffer passively from the increase in travel times. This leads to a pattern of flows $\mathbf{x}^{\text{SA}}$, representing the state of the system immediately after the disruption.

3) *Progressive Assignment*. This section of the model, shown in the Progressive Assignment algorithm box reported below, is essentially responsible for the evolution of the system and consists of three components. (1) The *Target Assignment* is responsible for allocating, at each iteration, the flows of each origin-destination pair through an N-path restricted User-Equilibrium traffic assignment model (Lin and Leong, 2014). The demand

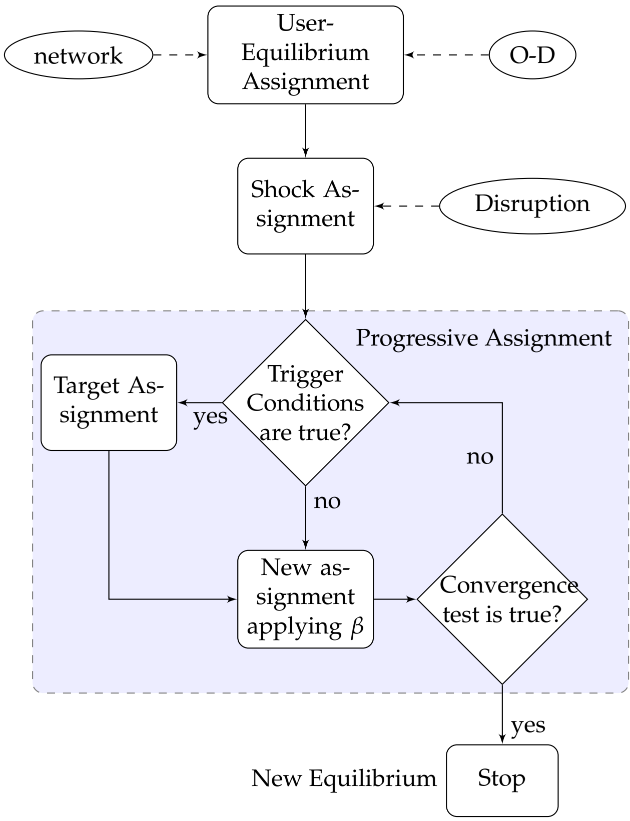



FIGURE 5.1: The assignment model.

flow of each pair is assigned considering only a sub set of all available paths $\mathcal{L}_h^n$. The evolution of this set for each OD pair depends on some Trigger Conditions. This leads to an "ideal" link flow pattern $\mathbf{z}(n)$ , given the current conditions, called *target*, representing the state towards which the system tends to move. (2) *Applying the inertia coefficient* $\beta$, the actual pattern of flows is obtained through the following dynamic equation:

$$\mathbf{x}(n) = \beta \mathbf{x}(n-1) + (1-\beta)\mathbf{z}(n) \tag{5.3}$$

where $\beta \in [0,1]$. (3) Finally, *The Trigger Conditions* are the most important element since they are responsible for managing the sets of paths from which each target assignments is derived, influencing the direction the system is evolving forward.

On each day, two conditions are evaluated: **condition 1** verifies if someone on the network is experiencing higher travel times than acceptable. i.e.

$$\exists \text{ a path } k \in \mathcal{L}_h^n : \frac{C_k^h(\mathbf{f}(n)) - C_k^h(\mathbf{f}^{\text{UE}})}{C_k^h(\mathbf{f}^{\text{UE}})} > \Omega \tag{5.4}$$

**condition 2** verfies if any origin-destination pair still has an unused path available, i.e.

$$|\mathcal{K}_h| > |L_h^n| \tag{5.5}$$

where (5.5) simply states that **condition 2** is met when the cardinality of the set of all available paths for OD $h$ is greater than the cardinality of the set of paths used for assignments up to day $n$.

Only for those origin-destination pairs for which **conditions 1** and **2** are true at the same time an unused path from set $\mathcal{K}_h$ is added to the set of currently available paths.

This leads to the set of paths $\mathcal{L}_h^n$, that can be used in the next target assignment. Only if **condition 1** and **2** are true at the same time for at least one OD pair it is useful to trigger the Target Assignment. If it is not the case in order to avoid a waste of computational resources the new target assignment can be set as equal to the previous one $\mathbf{z}(n) = \mathbf{z}(n-1)$.

**Progressive Assignment Algorithm**

---

**Input**: $\mathbf{x}^{\mathrm{SA}}$, $\mathbf{c}(\mathbf{x}^{\mathrm{SA}})$, $\mathbf{C}(\mathbf{f}^{\mathrm{UE}})$, $\mathcal{L}_h^{\mathrm{SA}}$
**Output**: $\mathbf{x}^{\mathrm{FE}}$

**1 Initialization**
$\mathbf{x}(0) = \mathbf{x}^{\mathrm{SA}}$, $\mathbf{c}(\mathbf{x}(0)) = \mathbf{c}(\mathbf{x}^{\mathrm{SA}})$, $\mathcal{L}_h^0 = \mathcal{L}_h^{\mathrm{SA}}$
*Checking Trigger Conditions for the first time*

$$\text{if } \exists \text{ a path } k \in \mathcal{L}_h^0 : \frac{C_k^h(\mathbf{f}(0)) - C_k^h(\mathbf{f}^{\mathrm{UE}})}{C_k^h(\mathbf{f}^{\mathrm{UE}})} > \Omega$$

proceed to the next step and set counter $n = 1$
**2 Target UE assignment**
Given for each pair $h$ the actual sets of paths $L_h^n$, perform a User Equilibrium traffic assignment by applying the Frank-Wolfe algorithm. This yields the target set of flows:

$$\mathbf{z}(n) = [z_1(n), z_2(n), \ldots, z_{\mathrm{A}}(n)]$$

**3 Applying the inertia coefficient $\beta$**:
The flows on links are computed as:

$$\mathbf{x}(n) = \beta\mathbf{x}(n-1) + (1-\beta)\mathbf{z}(n)$$

**4 Updating of travel costs**
update link travel costs: $\mathbf{c}(\mathbf{x}(n))$
update path travel costs: $\mathbf{C}(\mathbf{x}(n)) = \Delta' \mathbf{c}(\mathbf{x}(n))$

**5 Check Trigger Conditions**
Conditions 1 and 2 are checked. For only those OD pairs where both conditions are verified at the same time, add a path from $\mathcal{K}_h$ to the set of available paths at the next iterations. This yields a set of paths $L_h^n$. If conditions 1 and 2 are true at the same time for at least one OD pair after step (6) restart from step (2), otherwise after step (6) restart from (3) and the new target assignment will remain the one previously calculated:

$$\mathbf{z}(n+1) = \underline{z}(n)$$

**6 Convergence Test**
if a convergence test criterion is met, stop. The current solution is the new equilibrium:

$$\mathbf{x}^{\mathrm{FE}} = \mathbf{x}(n)$$

otherwise set $n = n + 1$ and go to step (2) or (3) depending on the outcome of step (5).

## 5.3 Performance metrics

Similar to what discussed by Omer et al. (2012), Faturechi and Miller-Hooks (2014) and Bhavathrathan and Patil (2015), to assess the level of system operativity, the following performance measures based on users' travel cost are proposed.

The global system performance is defined as follows:

$$P^n = r^n / r^{\mathrm{UE}} \tag{5.6}$$

where, $r^n$ and $r^{\mathrm{UE}}$ are the inverse of the total travel cost on day $n$ ($TTC^n$) and total travel cost in pre-disruption scenario ($TTC^{\mathrm{UE}}$) respectively. As a consequence, the performance during the evolution of the system is expressed as a percentage of the pre-disruption performance, that is, when the system was operating under normal conditions.

Similarly, it is possible to define the quality of the network as perceived by the users of each OD pair as follows:

$$P_h^n = r_h^n / r_h^{\mathrm{UE}} \tag{5.7}$$

where similarly with what has been defined for the overall system, $r_h^n$ and $r_h^{\mathrm{UE}}$ are the inverse of the average travel cost $\bar{C}_h^n$ experienced by users of OD pair $h$ on day $n$ and average travel cost in a pre-disruption scenario $\bar{C}_h^{\mathrm{UE}}$ experienced by the same users respectively.

## 5.4 Implementation and Results

The progressive assignment model is evaluated on the Nguyen-Dupuis test network (Nguyen and Dupuis, 1984). This network is represented by an oriented graph consisting of 13 nodes and 19 links. The transport demand vector is as follows:

- $q_{12} = 50$
- $q_{13} = 10$
- $q_{42} = 40$

- $q_{43} = 20$.

The performance functions specific to each link $a \in \mathcal{A}$ are assumed linear $t_a(x_a) = \omega_a x_a + \theta_a$, where $\omega_a$ is the marginal travel cost while $\theta_a$ is the free flow travel cost. They are both strictly greater than zero in order to satisfy the condition given travel cost function.

Both the system's User-Equilibrium before the disruption and the Target Assignments performed on each day are solved using the convex combination algorithm originally suggested by Frank and Wolfe (1956). Figure 5.2 and table 5.2 show the paths used by each OD pair at the equilibrium and the network with the traffic flows assigned to each link before the occurrence of the disruption respectively.

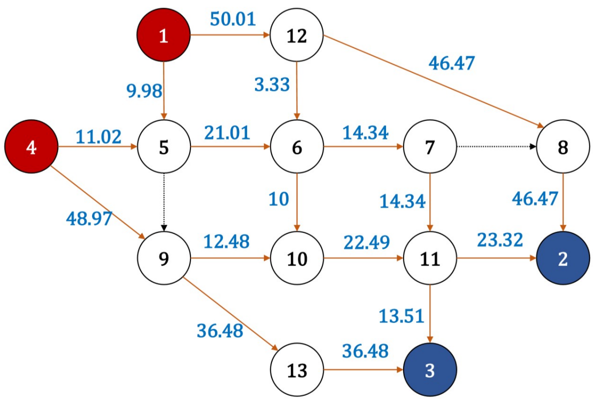



FIGURE 5.2: Nguyen-Dupuis network and pre-disruption assignment.

TABLE 5.2: Paths used at the equilibrium.

| O-D pair | paths |
|---|---|
| 1-2 | [1, 12, 8, 2], [1, 12, 6, 7, 11, 2] |
| 1-3 | [1, 5, 6, 10, 11, 3] |
| 4-2 | [4, 9, 10, 11, 2], [4, 5, 6, 7, 11, 2] |
| 4-3 | [4, 9, 13, 3], [4, 9, 10, 11, 3] |

Once the system state is determined under normal conditions, the disruption is obtained by removing the link between nodes 12 and 8. The link has been chosen as peripheral as possible, in order to emphasize any propagation phenomenon in the performance deterioration and to avoid a dynamics excessively fast to be appreciated. Table 5.3 shows the evolution of the system, from the initial condition (marked as EU), through the shock caused by the link cancellation (marked as SA) to the reaching of the new equilibrium by applying the Progressive Assignment procedure for $n = 1, \ldots, 17$.

TABLE 5.3: System evolution for $\Omega = 0.2$, $\beta = 0.6$.

| Iteration | $TTC^n$ | $\bar{C}^n_{12}$ | $\bar{C}^n_{13}$ | $\bar{C}^n_{42}$ | $\bar{C}^n_{43}$ |
|---|---|---|---|---|---|
| UE | 19 963.47 | 196.04 | 131.49 | 186.41 | 127.95 |
| SA | 41 750.75 | 525.08 | 326.59 | 397.03 | 150.71 |
| 1 | 42 376.13 | 532.86 | 332.97 | 390.48 | 154.55 |
| 2 | 40 310.74 | 476.68 | 302.78 | 394.94 | 159.48 |
| 3 | 31 959.53 | 362.99 | 252.76 | 286.53 | 152.19 |
| 4 | 28 833.24 | 338.20 | 238.85 | 267.92 | 147.81 |
| 5 | 27 331.00 | 314.51 | 216.32 | 254.67 | 145.78 |
| 6 | 26 434.39 | 289.85 | 195.71 | 244.79 | 144.72 |
| 7 | 26 096.41 | 281.68 | 192.15 | 238.87 | 144.07 |
| 8 | 25 966.69 | 278.48 | 188.43 | 235.32 | 143.69 |
| 9 | 25 915.17 | 276.55 | 187.79 | 233.19 | 143.46 |
| 10 | 25 893.68 | 275.40 | 187.41 | 231.64 | 143.32 |
| 11 | 25 884.21 | 274.71 | 187.18 | 230.99 | 143.24 |
| 12 | 25 879.76 | 274.29 | 187.04 | 230.61 | 143.19 |
| 13 | 25 877.53 | 274.04 | 186.96 | 230.37 | 143.16 |
| 14 | 25 876.35 | 273.89 | 186.91 | 230.23 | 143.14 |
| 15 | 25 875.69 | 273.81 | 186.88 | 230.15 | 143.13 |
| 16 | 25 875.11 | 273.75 | 186.86 | 230.10 | 143.12 |
| 17 | 25 875.02 | 273.71 | 186.83 | 230.08 | 143.08 |

The *user tolerance index* $\Omega$ has been set to 0.2 and the *inertia coefficient* $\beta$ to 0.6. This means that users are insensitive to increases in travel times of less than 20% while at each iteration 60% of the flows of the current assignment are affected by the previous one. In the first column of table 5.3 the Total Travel Cost $TTC^n$ for the whole system is shown and it is computed as follows.

$$TTC^n = \sum_{a \in \mathcal{A}} x_a(n) \cdot c_a(x_a(n)) \tag{5.8}$$

In the other columns the average OD travel cost is reported.

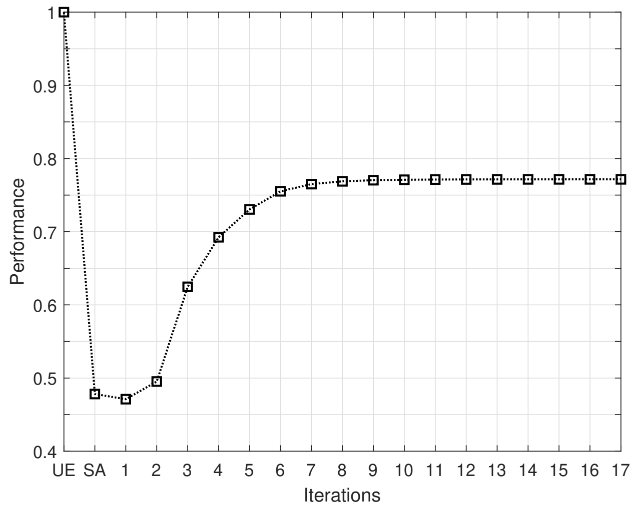



FIGURE 5.3: Global system performance ($\Omega = 0.2$, $\beta = 0.6$).

Fig. 5.3 shows the evolution of system performance as defined in (5.6) and corresponding to $TTC^n$ values given in table 5.3. As it can be seen, immediately after the disruption, the performance of the system deteriorates dramatically. Total Travel Cost goes from a value of about 19 963 to a value of 41 750, approximately 47% of the initial performance, thus recording a deterioration in overall performance of more than 50%. After this spike, as the flows of each OD pair are progressively reassigned over a larger set of paths, the performance of the system gradually improves. At approximately the 8th iteration, the new equilibrium is reached settling around 77% of the initial performance, resulting in a definitive performance loss of about 23%. This means that, if we consider the fact that $\Omega$ has been set to 0.2, definitely some users remain unsatisfied even at the new equilibrium, yet they are not able to do any better because of the new network topology.

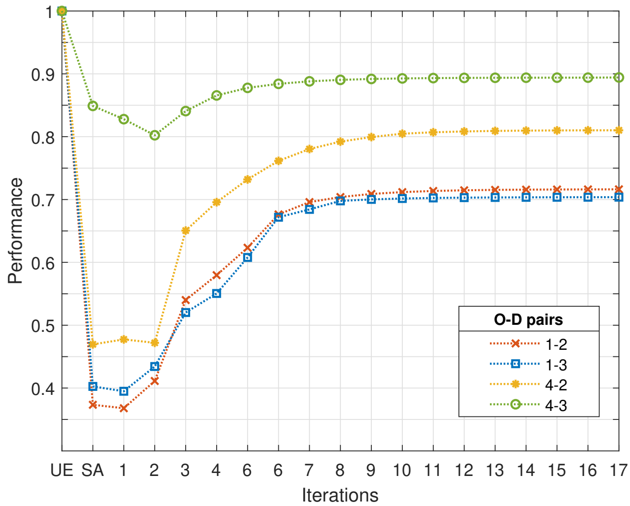



FIGURE 5.4: Performance on paths ($\Omega = 0.2$, $\beta = 0.6$).

Considering instead the metrics defined in (5.7), fig. 5.4 shows the evolution of the performances as they are experienced by the users of each OD pair traveling on their respective paths. In this case a cascading effect in the deterioration of the performance can be noticed which, starting from the area directly involved in the cancellation of the link, expands with less magnitude as it involves the remaining parts of the network. Specifically, the more the paths of a particular OD pair are influenced by the link omission, the more they are affected by the disturbance. Looking at table 5.2, it can be noticed that $1-2$ is the only one of the origin-destination pairs to be directly involved in the disruption and for this reason it is the one that suffers the most. Path $[1, 12, 8, 2]$ used by the users of the $1-2$ pair at equilibrium is no longer available and as a consequence this transport demand spreads, during the period following the disruption, on other portion of the network influencing other users. Among all, the $4-3$ pair is the least affected by the disturbance, showing a maximum deterioration in performance on its paths of about 20%. This is consistent with the fact that the users of this pair use paths that are not closely connected with those used by the users of the other pairs, especially those of the $1-2$ pair.

In the following, fig. 5.5 and fig. 5.6 show how the evolution of the system is influenced respectively by the inertia coefficient $\beta$ and user tolerance index $\Omega$.

Consistently with expectations, as shown in fig. 5.5 the inertia coefficient $\beta$ influences the speed at which the system converges to the new equilibrium. The higher the number of users willing to use alternative paths, the faster the system evolves towards a new stable state. By contrast, the variation of $\beta$ has no influence on determining what the value of this new equilibrium will be, except in the extreme case of $\beta = 1$. In this case, the state of the system of the $n+1$-iteration does not actually evolve further, once the disruption has occurred.

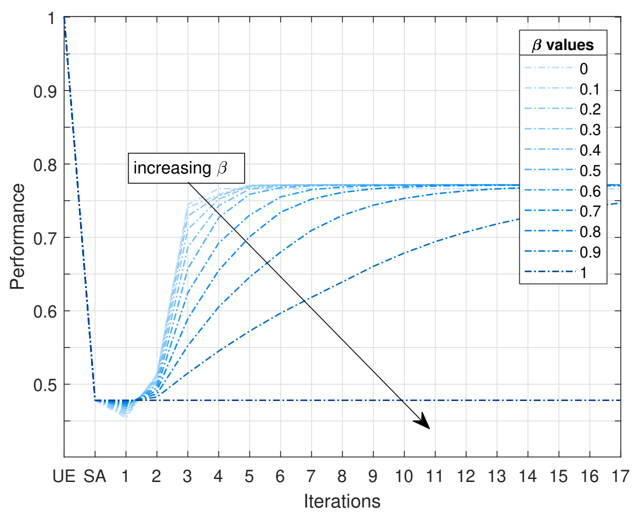



FIGURE 5.5: Impact of $\beta$ on system performance evolution ($\Omega = 0.2$).

Fig. 5.6 shows how the global system performance is influenced by the user tolerance index $\Omega$. The response of the system is evaluated for values of $\Omega$ ranging from 0 to 2. More specifically, starting from a scenario in which users do not tolerate any increase, however small, in travel costs ($\Omega = 0$) to one in which they are insensitive to increases in travel costs lower than 200% of pre-disruption values ($\Omega = 2$), we evaluated the response of the system by applying values of the coefficient $\Omega$ obtained by discretizing the interval each 0.2 units. As expected, the values on which performance stabilizes, once the new equilibrium is reached, are partly influenced by user preferences.

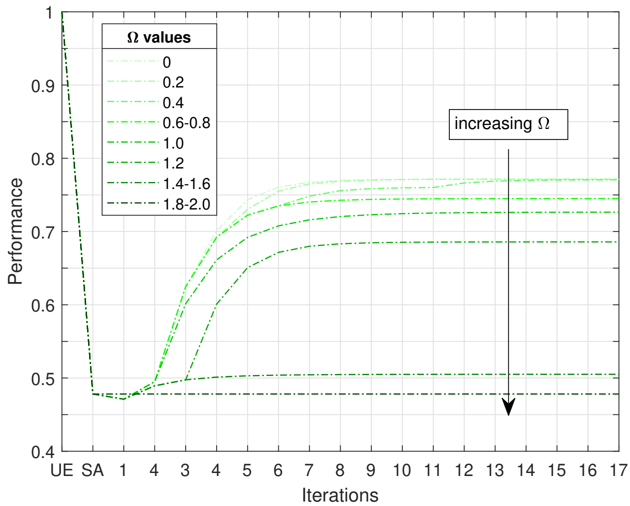



FIGURE 5.6: Impact of Ω on system performance evolution ($\beta = 0.6$).

In detail, the more we increase the tolerance of users to increases in travel times the higher the new equilibrium will be. In other words, highly tolerant users have less incentive to use new paths to improve their travel times.

It is worth noting that, the nature of the relationship between the level that the new equilibrium will reach and the values of Ω is strongly discontinuous. Even with major increases in the coefficient, the final equilibrium achieved by the system may not change. On the contrary, sometimes it can happen that for small variations of Ω the new equilibrium changes considerably. This is a consequence of how the model is designed. The user tolerance coefficient affects the set of paths on which the flows of users can be loaded at each next iteration. As long as the travel costs experienced by the users do not imply a percentage increase of more than Ω, the set of paths on which the assignment will be performed will not change regardless of Ω and as a consequence, not even the dynamics of the system. However, when $\Omega$ has varied enough to match this critical value for at least the users of one OD pair, the set of paths of the same ones changes. As a result, these traffic flow will spread over a larger portion of the network, changing the travel costs of other users, eventually causing a cascading phenomenon that results in a large traffic flow re-assignment. This is clearly visible in fig. 5.6 if we look at the performance trends for $\Omega = 1.4$ and $\Omega = 1.2$.

Finally, it is interesting to note that even if we set $\Omega = 0$, there is still a considerable gap between the original performance of the system and that given by the new equilibrium. This is due to the fact that, as much as users strive, the new network topology does not allow them to get better conditions than the original ones. For this reason, we can conclude that according to the model part of the performance differences between the two equilibriums, the one before and the one reached after the disruption, can be influenced by the users and their preferences while the remaining part is determined exclusively by the network and by the location where the disruption takes place.

## 5.5 Conclusions

In this chapter an assignment model capable of representing the evolution of a traffic network in the short term after the occurrence of a critical event has been presented. The Progressive Assignment, taking into account the users tolerance to increases in travel costs and the intrinsic inertia of the system, controls the sets of paths on which the flow will be assigned on each day. From the results presented, the model appears to be able to represent some aspects of the evolution of the system in a reasonable way.

# Chapter 6

# Two-Stage Multi-Class Modeling Approach for Multi-Modal Transport Networks

Based on *Pasquale et al. (2022)*

A particularly significant perturbation affecting one transportation subsystem, such as the road or rail network, can cause cascading phenomena affecting other transportation modes as well. In this regard, a multimodal scenario analysis able to represent flow shifts across different transport modes is introduced in this chapter. The need to move passengers and freight is becoming increasingly pervasive in conducting daily human activities and more generally in the social and economic advancement of countries. As a result, the availability and accessibility of transport infrastructures and mobility services allow this need to be met and, at the same time, increase the attractiveness of an area. In this regard, it important to quantify the capacity to satisfy the transport demand of a geographical area and to try to predict how the transport systems included in it will react to unforeseen events. These latter may appear in different forms as changes in demand, e.g. the transfer of passengers from public to private transport that have occurred due to the ongoing pandemic, or modification of the infrastructure network layout due to the verification of disruptive events.

Model-based approaches are the foremost methodologies for evaluating the performance of transportation networks because they allow:

- to quantify the efficiency of transport networks subject to different scenarios through the calculation of performance indexes (e.g., average travel times, fuel consumption, pollutant emissions, etc.);
- to evaluate the effects produced by the occurrence of critical events (e.g., infrastructure collapse, natural disasters limiting the functionality of transport systems, terrorist attacks, etc.);
- to evaluate the effects produced by the introduction of new systems (e.g., new infrastructure, modification of existing layout, etc.);
- to develop or test regulation policies.

In the transport network application field, models can be generally distinguished into assignment models and simulative models. Assignment models are adopted to represent the route choices of users and to evaluate the distribution of flows on a network. Simulative models are used to represent the impact of route choices on transport networks by describing the flow evolution over time. By analyzing the literature on these topics, it is possible to observe that most of the researches' efforts have been focused so far on the definition of assignment and simulative models on mono-modal transport networks, i.e. road networks or railway networks only (Cascetta, 2009; Hoogendoorn and Bovy, 2001; Gille, Klemenz, and Siefer, 2008). However, if the goal is to analyze the accessibility of an area in its wholeness or to give better information to users, it is necessary to adopt more extensive models, i.e. multi-modal models, that simultaneously consider different modes of transports and the possibility of transfer among them.

Another aspect that cannot be overlooked in the analysis of transport networks concerns the interdependence between different transport modalities. Interdependence among transportation modes can be expressed in several ways. It can be related to the pursuit of the transport activity itself, in this case several transport services operate synergistically to allow the satisfaction of the mobility demand of users and the efficient distribution of goods and services. Interdependence can be related to physical reasons such as the overlap of different routes, even belonging to different transport modes, through bridges or tunnels, or there can be hidden relationships, such as those analyzed by Chen et al. (2009). Regardless of how these interdependencies occur, they imply an increased vulnerability of the transport network as a whole. In fact, it is possible that critical events affecting even one mode of transport could cause a ripple effect involving other transport modalities or, in the worst case, the entire transport network. For these reasons, the development of multi-modal models represents an important step in the analysis of complex and highly interdependent systems such as transport networks.

This work falls in this field of research by proposing a two-stage model in which multi-modal assignment and multi-modal simulation are combined. More in detail, the proposed modeling scheme, based on preliminary works (Pasquale et al., 2021a; Pasquale et al., 2021b), is suitable to represent large-scale transportation networks in which the considered transport modes are road and rail connected to each other through appropriate intermodal connections. Another peculiarity of this modeling framework

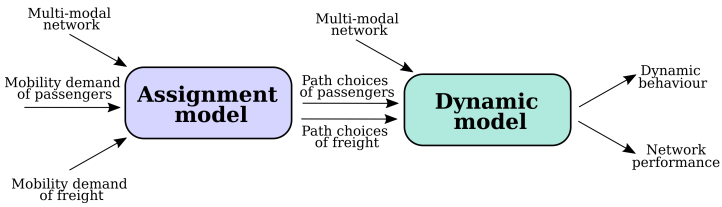



FIGURE 6.1: Sketch of the two-stage modeling framework.

is that the user demand is multi-class, hence it is distinguished into passengers and cargo units (e.g., containers). This distinction is of particular importance because passenger and freight flows may be characterized by different behaviors and may be subject to different restrictions, such as on route choices. For these reasons, both the assignment and the dynamic model are multi-class. Note that, compared with Pasquale et al. (2021a) and Pasquale et al. (2021b), in the present work the assignment model is multi-class (while in the previous versions there was no distinction of user classes) and the dynamic model is improved.

To summarize, as sketched in Fig. 6.1, the modeling framework proposed in this work is composed of two stages:

- *a multi-class multi-modal assignment model*;
- *a multi-class multi-modal discrete-time dynamic model*.

This chapter concludes with the application of the proposed modeling framework on a benchmark network. The focus of this application is to show a possible use of the proposed framework by simulating the loss of operability of one connection of the multi-modal network and analyzing the ripple effect on the rest of the network. Other possible applications of the proposed methodology are discussed at the end.

The present chapter is organized as follows. A state of the art in multi-modal transport network models is presented in section 6.1. In section 6.2 the general features of the proposed modeling framework and the basic notation are introduced. In section 6.3 the proposed assignment model is outlined, while the macroscopic discrete-time dynamic model is described in section 6.4. The application of the proposed methodology to a test network is shown in section 6.5, while section 6.6 discusses possible uses of the proposed approach. Finally some conclusive remarks are gathered in section 6.7.

## 6.1 State of the art on multi-modal transport networks models

The literature on multi-modal network models is much less established than the one considering single modes of transport. The majority of the works addressing multi-modal transport networks concerns the definition of multi-modal assignment models mostly referring to urban areas. Indeed, multi-modal traffic assignment has been studied for some decades (Dafermos, 1976; Florian, 1977; Florian and Nguyen, 1978; Sheffi and Daganzo, 1978), by considering the so-called hyper-networks, which allow to include in the network both links representing a portion of the real multi-modal transport network (road, rail, private or public transport) and links associated with users' decisions.

Multi-modal traffic assignment approaches have also been studied more recently, e.g. Lo, Yip, and Wan (2003) represented the multi-modal transport network of an urban area as a graph composed of a number of sub-networks, each of which is associated with a transportation mode. Then the multi-modal network is translated into an augmented-state multi-modal network in which transfer rules and transfer probabilities between the different transport modes are defined. In that network, multi-modal routes are determined with standard assignment models. Fu, Lam, and Chen (2014) proposed a stochastic model in order to account for the reliability of the chosen transport modes and paths, while Pi, Ma, and Qian (2019) considered different modes, only-driving, carpooling, ride-hailing, public transit and park-and-ride, for a urban transportation system including private cars, freight trucks, buses, and so on.

Other works involve the definition of dynamic programming problems aimed at identifying multi-modal routes or at regulating the transition between different types of transport. In the work by Zhang et al. (2011), the transportation modes are distinguished into private and public modes of transports and connected with some abstract links that allow the transfer between the considered modes. The only attributes associated with the elements of the network are the travel times that are estimated considering the length and speed of a link and possible waiting times. Then, a rule-based algorithm is proposed to suggest travel information. An application based on learning techniques is instead presented by Liu et al. (2020). In this work, data from a large-scale multi-modal network are analyzed and both spatio-temporal and semantic coherence relationships are identified as the basis for route choice. Based on these findings, a hierarchical multitask learning module is developed to represent different transportation modes and to guide route choice.

Several simulative-based approaches have been presented (Loder et al., 2017; Liu and Geroliminis, 2017; Elbery et al., 2018; Bucchiarone, De Sanctis, and Bencomo, 2020). Both approaches presented by Loder et al. (2017) and Liu and Geroliminis (2017) involve the use of the macroscopic fundamental diagram (MFD). Specifically, Loder et al. (2017) consider three transport modalities, i.e. private vehicle transport, public vehicle transport and pedestrian transport. The three transport modes are represented through a three-dimensional MFD that allows to evaluate the accumulation of private and public vehicles according to different traffic scenarios. In the work by Liu and Geroliminis (2017), the considered transport modes are private vehicle transport and public road transport. Private transportation is represented through the MFD, while constant transfer speeds are considered for public transportation. The derived multi-modal model is used for the definition of adaptive pricing policies. Finally, Elbery et al. (2018) and Bucchiarone, De Sanctis, and Bencomo (2020) propose agent-based approaches. Specifically, the model presented by Elbery et al. (2018) introduces an agent-based framework to combine different transportation systems, namely vehicular, bicycle, pedestrian and railway, which in turn are represented with different simulators (INTEGRATION which is a meso-scopic simulator based on the car-following model, BPSim that is a bike and pedestrian commercial simulator and RailSim which is a commercial railway simulator). Bucchiarone, De Sanctis, and Bencomo (2020) propose an agent-based simulation to test innovative planning methods involving pedestrians and autonomous shuttles in large urban areas.

Note that in none of these approaches the assignment models are combined with simulative models. Furthermore, in most of these works the dynamics of the multi-modal system is not explicitly represented, and therefore travel times on multi-modal transport networks are either fixed a priori or provided by other simulation models. Another innovative aspect of the work here proposed, which was not found in the literature analysis, is the explicit representation of multiple classes of users. Thus, to the best of our knowledge, this is the first work in which multi-class multi-modal assignment and simulation models are combined.

## 6.2 General features and basic notation

This section provides the main features and the principal notation used in the proposed modeling framework. As shown in Fig. 6.1, this scheme consists of two stages: an assignment model defined to allocate the demand of passengers and freight on a multi-modal transport network, and a discrete-time dynamic model that allows to replicate the evolution of the system over time. It is worth noting that the discrete-time dynamic model receives as input the mobility demand and the path choices (route and modal choices) defined through the multi-modal assignment model.

Both the assignment model and the dynamic model are based on a regional multi-modal transport network in which the considered transport modes are road transport, represented by a highway network, and rail transport. The two modes of transport are connected with some intermodal arcs distinguished depending on the flow class, i.e. passengers or cargo units, which they can receive. The multi-modal transport network is represented by means of an oriented graph, as depicted in Fig. 6.2, denoted with $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, in which $\mathcal{N}$ indicates the set of nodes, whereas $\mathcal{A} = \mathcal{A}^{\mathrm{H}} \cup \mathcal{A}^{\mathrm{R}} \cup \mathcal{A}^{\mathrm{Ip}} \cup \mathcal{A}^{\mathrm{If}}$ represents the set of arcs. Each subset of arcs is defined as follows:

- $\mathcal{A}^{\mathrm{H}}$ is the set of highway connections;

- $\mathcal{A}^{\mathrm{R}}$ is the set of railway connections;
- $\mathcal{A}^{\mathrm{Ip}}$ is the set of intermodal arcs for passengers;
- $\mathcal{A}^{\mathrm{If}}$ is the set of intermodal arcs for freight.

Note that intermodal arcs are distinguished for passengers and for freight because modal shifts for passengers and freight typically occur at different locations and in different ways.

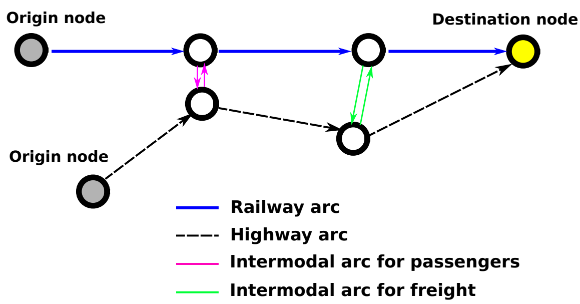



FIGURE 6.2: Sketch of the multi-modal transport network.

Moreover, let $P(i)$ indicate the set of nodes preceding node $i$ and $S(i)$ the set of nodes succeeding node $i$, $i \in \mathcal{N}$. Let $\Delta_{i,j}$ [km] denote the length of each arc $(i,j) \in \mathcal{A}^{\mathrm{H}} \cup \mathcal{A}^{\mathrm{R}}$.

The network is defined in an origin-destination-oriented mode in which $J^{\mathrm{O}} \subseteq \mathcal{N}$ represents the set of all possible origin nodes, $J^{\mathrm{D}} \subseteq \mathcal{N}$ represents the set of all possible destination nodes. The nodes which are not origins nor destinations simply allow the transit between successive arcs. The intermodal arcs, for both flow classes, are considered as fictitious arcs that allow modal transfers between road transport and rail transport and vice versa. For this reason, an origin node $o \in J^{\mathrm{O}}$ cannot be followed only by one or more intermodal arcs, and similarly, a destination node $d \in J^{\mathrm{D}}$ cannot be preceded only by one or more intermodal arcs.

In the proposed approach the flows of users are distinguished in passengers and freight, with superscript $c$ denoting the class of users. In particular we define with $c$ the class of users. In particular we define with $c = 1$ the passenger flow and with $c = 2$ the freight flow. Then, $D^{od,c}$ indicates the demand of class $c$ that originates at node $o \in J^{\mathrm{O}}$ and has destination at node $d \in J^{\mathrm{D}}$: this is the total demand associated with the *od* pair for the whole simulation horizon. This demand defined for all *od* pairs represents the so-called "origin-destination matrix". Let us also specify that the passenger demand, i.e. the demand for class $c = 1$, is expressed in number of passengers, while the demand of freight, corresponding to $c = 2$, is expressed in number of cargo units.

The main parameters and variables of the proposed approach are summarized in Tables 6.1-6.4. Specifically, Table 6.1 collects all the parameters that are used in both the assignment model and the dynamic simulation model. Table 6.2 and Table 6.3 collect the parameters and variables used in the assignment model only respectively, while Table 6.4 reports parameters and variables used in the dynamic simulation model only.

TABLE 6.1: Parameters common to the two models in the two-stage modeling framework

| Parameter | Description |
|---|---|
| $\Delta_{i,j}$ | Arc length [km] for all $(i,j) \in \mathcal{A}^{\mathrm{H}} \cup \mathcal{A}^{\mathrm{R}}$ |
| $v^{\mathrm{H}}_{i,j}$ | Maximum freeway speed in [km/h] for all $(i,j) \in \mathcal{A}^{\mathrm{H}}$ |
| $\omega_{i,j}$ | Congestion wave speed in freeway in [km/h] for all $(i,j) \in \mathcal{A}^{\mathrm{H}}$ |
| $n^{\max}_{i,j}$ | Maximum number of vehicles for all $(i,j) \in \mathcal{A}^{\mathrm{H}}$ |
| $v^{\mathrm{R}}_{i,j}$ | Maximum railway speed in [km/h] for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $h_{i,j}$ | Average time headway in [h] for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $s^{\min}_{i,j}$ | Minimum average space headway in [km] for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $C^{\mathrm{p}}$ | Capacity of a passenger train in number of passengers for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $C^{\mathrm{f}}$ | Capacity of a freight train in rail wagons for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $L$ | Length of a train in [km] for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $\alpha_{i,j}$ | Number of time steps required to cross an intermodal passenger arc $(i,j) \in \mathcal{A}^{\mathrm{Ip}}$ |
| $\gamma_{i,j}$ | Number of time steps required to cross an intermodal freight arc $(i,j) \in \mathcal{A}^{\mathrm{If}}$ |
| $\eta$ | Average number of passengers per vehicle |
| $\varsigma$ | Conversion factor used to convert trucks into passenger car equivalent |

As it will be described in section 6.4, the dynamic behavior of the whole network is represented with a discrete-time model in which the time horizon is divided in $K$ time steps, where $k = 1, \ldots, K$ indicates the temporal stage, and $T$ [h] represents the sample time interval. In order to ensure a correct time discretization, the length of the time interval $T$ must allow a proper dynamic evolution of the system, therefore the length of the time step is chosen in order to verify, for any arc $(i,j) \in \mathcal{A}^{\mathrm{H}} \cup \mathcal{A}^{\mathrm{R}}$, the following condition:

$$T \leq \min\left\{\frac{\Delta_{i,j}}{v^{\mathrm{H}}_{i,j}}, \frac{\Delta_{i,j}}{v^{\mathrm{R}}_{i,j}}\right\} \tag{6.1}$$

TABLE 6.2: Parameters present only in the assignment model described in section 6.3

| Parameter | Description |
|---|---|
| $\delta^{od,1}_{i,j,l}$ | Binary parameter that defines the belonging of an arc $(i,j) \in \mathcal{A}$ to a passenger path associated to the *od* pair |
| $\delta^{od,2}_{i,j,l}$ | Binary parameter that defines the belonging of an arc $(i,j) \in \mathcal{A}$ to a freight path associated to the *od* pair |
| $C^{\mathrm{time}}_{i,j}$ | Cost per time unit used to estimate arc-traveling costs for freight flows for all arc $(i,j) \in \mathcal{A}$ |
| $C^{\mathrm{space}}_{i,j}$ | Cost per space unit used to estimate arc-traveling costs for freight flows for all arc $(i,j) \in \mathcal{A}$ |
| $C^{\mathrm{fix}}_{i,j}$ | Fixed cost used to estimate arc-traveling costs for freight flows for all arc $(i,j) \in \mathcal{A}$ |
| $\bar{t}^{1}_{i,j}$ | Average travel time over arc $(i,j) \in \mathcal{A}$ for passengers |
| $\phi$ | Parameter used to define a performance function for passengers |
| $M$ | Large coefficient chosen arbitrarily |

TABLE 6.3: Variables present only in the assignment model described in section 6.3

| Variable | Description |
|---|---|
| $f^{od,1}_{l}$ | Decision variable related to passenger flow associated with the *od* pair using a specific path $l$ |
| $f^{od,2}_{l}$ | Decision variable related to freight flow associated with the *od* pair using a specific path $l$ |
| $x^{1}_{i,j}$ | Decision variable related to the total passenger flow on arc $(i,j) \in \mathcal{A}$ |
| $x^{1,\mathbf{UE}}_{i,j}$ | Decision variable related to the total passenger flow on arc $(i,j) \in \mathcal{A}$ satisfying User-Equilibrium conditions |
| $x^{2}_{i,j}$ | Decision variable related to the total freight flow on arc $(i,j) \in \mathcal{A}$ |
| $\tau^{1}_{i,j}(\cdot)$ | Performance functions for passengers of arc $(i,j) \in \mathcal{A}$ |
| $\tau^{2}_{i,j}(\cdot)$ | Performance functions for freight of arc $(i,j) \in \mathcal{A}$ |
| $c_{i,j}(\cdot)$ | Total travel costs for freight on the arc $(i,j) \in \mathcal{A}$ |

TABLE 6.4: Parameters present only in the dynamic model described in section 6.4

| Parameter | Description |
|---|---|
| $N_{i,j}^{\max}$ | Maximum number of trains for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $\epsilon_{i,j}^{c}$ | Conversion factor defined for each class $c$ and for each arc $(i,j) \in \mathcal{A}$ to model the transition from a railway arc to a road arc and vice versa |
| $\xi_{i,j}^{c}$ | Conversion factor defined for each class $c$ and for each arc $(i,j) \in \mathcal{A}$ to correctly computing the flow entering from an origin node |

TABLE 6.5: Variables present only in the dynamic model described in section 6.4

| Variable | Description |
|---|---|
| $n_{i,j}^{od,c}(k)$ | Number of units of class $c$ in arc $(i,j) \in \mathcal{A}$, associated with the $od$ pair at time step $k$ |
| $n_{i,j}^{\mathrm{tot}}(k)$ | Total number of vehicles on arc $(i,j) \in \mathcal{A}^{\mathrm{H}}$ |
| $N_{i,j}^{\mathrm{tot}}(k)$ | Total number of trains in arc $(i,j) \in \mathcal{A}^{\mathrm{R}}$ |
| $I_{i,j}^{od,c}(k)$ | Number of units of class $c$ entering arc $(i,j) \in \mathcal{A}$, associated with the $od$ pair at time step $k$ |
| $O_{i,j}^{od,c}(k)$ | Number of units of class $c$ exiting arc $(i,j) \in \mathcal{A}$ associated with the $od$ pair at time step $k$ |
| $w_{i,j}^{od,c}(k)$ | Number of units of class $c$ that would like to enter arc $(i,j) \in \mathcal{A}$ associated with the $od$ pair at time step $k$ |
| $W_{i,j}^{\mathrm{tot}}(k)$ | Total number of units that would like to to enter arc $(i,j) \in \mathcal{A}$ at time step $k$ |
| $\pi_{i,j}(k)$ | Surplus rate of units that cannot enter at time step $k$ in the arc $(i,j) \in \mathcal{A}$ |
| $S_{i,j}^{od,c}(k)$ | Number of units of class $c$ that would like to exit arc $(i,j) \in \mathcal{A}$, associated with the $od$ pair at time step $k$ |
| $\beta_{i,j}^{od,c}(k)$ | Splitting rates of class $c$ in arc $(i,j) \in \mathcal{A}$, associated with the $od$ pair at time step $k$ |
| $q^{od,c}(k)$ | Number of units of class $c$ associated with the $od$ pair, that can actually enter the network from node $o \in \mathcal{J}^{\mathrm{O}}$at time step $k$ |
| $l^{od,c}(k)$ | Queue length of class $c$, associated with the $od$ pair, which waits at the origin node $o \in \mathcal{J}^{\mathrm{O}}$ at time step $k$ |
| $q_{i,j}^{\mathrm{res}}(k)$ | Residual capacity of arc $(i,j) \in \mathcal{A}$ at time step $k$ |
| $t_{i,j}(k)$ | Transfer time required to cover an arc $(i,j) \in \mathcal{A}$ at time step $k$ |
| $V_{i,j}(\cdot)$ | Steady-state speed relationships defined for railway and roadway arcs $(i,j) \in \mathcal{A}^{\mathrm{H}} \cup \mathcal{A}^{\mathrm{R}}$ for each time step $k$ |

## 6.3 The multi-modal assignment approach for passenger and freight flows

The purpose of the intermodal assignment model is to represent the spontaneous decisions of users about their routes. Before providing the details of the proposed approach, it is important to emphasize that the assignment model is a static approach in which all the mobility demand, estimated for a specific time window, is assumed to use the transport network simultaneously. As shown in Fig. 6.1, the result of this model is the redistribution of flows on the paths, i.e., $f_l^{od,1}$ for passengers and $f_l^{od,2}$ for freight, which are obtained by the solution of some optimization problems.

The flows thus obtained will be suitably transformed into splitting rates, see section 6.4, and used as input data for the dynamic model. The passenger and freight assignment models are treated separately, firstly introducing the multi-modal assignment model for passengers and, then, applying the multi-modal assignment model for freight, since the latter uses the results of the passenger assignment to define the freight route choices. In particular, for the passenger assignment procedure we assume that the marginal impact on the network due to the presence of freight is sufficiently small so that it does not decisively influence the behavior of passengers. This assumption is quite reasonable since, referring to the overall flows on a transport network, the freight component typically constitutes a rather low percentage compared with the flow of passengers.

To summarize, the assignment procedure presented below is conducted in an iterative manner in which at the first step the multi-modal assignment model for passenger flows described in section 6.3.1 is run. Then the results of this assignment are used to estimate the average passenger travel times on the network using the dynamic model described in section 6.4. Finally, the multi-modal freight assignment model, presented in section 6.3.2, is run using the average travel times defined in the previous step.

### 6.3.1 The multi-modal assignment model for passenger flows

Modeling the behavior and therefore the mobility choices of users is a challenging problem for which several approaches have been developed by researchers. One of the most accepted, and widely adopted, methodologies concerns the use of traffic assignment models. Given an origin-destination matrix, representing the users' demand, and knowing the functional characteristics of the network infrastructure, a traffic assignment model allows to estimate how users will be dispersed in the network, taking into account the relation between the infrastructure supply and the interaction of users mutual choices. The criteria underlying the mathematical representation of such behaviors have been covered in section 2.5.

In the present work, the N-Path Restricted User-Equilibrium traffic assignment model has been applied. Generally speaking, the User Equilibrium traffic assignment model aims to estimate the network equilibrium state such that no user has unilaterally any interest in choosing an alternative path, since no other path would guarantee lower travel times. Such equilibrium is achieved if Wardrop's first principle (Wardrop, 1952) is met, whereby "users choose the path that at a given time minimizes their own travel time". A way to compute the arc flows corresponding to such equilibrium involves finding the optimum solution of the Beckmann's Transformation (Beckmann, McGuire, and Winsten, 1956). The N-Path Restricted variant of the aforementioned model (Lin and Leong, 2014) is obtained by constraining the assignment to a priori defined sets of admissible paths, which do not necessarily include all the possible ones for each origin-destination pair. This allows to avoid all the paths that are theoretically possible but quite implausible in practice. In this work, all the paths involving more than one modal shift have been excluded. Then, denoted with $l$, $l = 1, \ldots, L$, a generic path existing in the network and with $\mathcal{L}^{od}$ the set of all possible paths connecting the $od$ pair, let $\mathcal{P}^{od} \subseteq \mathcal{L}^{od}$ be the set of admissible paths from origin node $o$ to destination node $d$.

The resulting optimization problem for the traffic assignment of passengers demand in the multi-modal transport network is the following.

**Problem 5.**

$$min\ z(\boldsymbol{x}) = \sum_{(i,j)\in\mathcal{A}} \int_0^{x^1_{i,j}} \tau^1_{i,j}(\omega)d\omega \tag{6.2}$$

*subject to*

$$\sum_{l \in \mathcal{P}^{od}} f_l^{od,1} = D^{od,1} \qquad o \in J^O, d \in J^D \tag{6.3}$$

$$f_l^{od,1} \geq 0 \qquad o \in J^O, d \in J^D, l \in P^{od} \tag{6.4}$$

$$x_{i,j}^1 = \sum_{o \in J^O} \sum_{d \in J^D} \sum_{l \in \mathcal{P}^{od}} f_l^{od,1} \cdot \delta_{i,j,l}^{od,1} \qquad (i,j) \in \mathcal{A} \tag{6.5}$$

where $f_l^{od,1}$ and $x_{i,j}^1$ are, respectively, the flow of class 1 of $od$ pair using path $l$ and the total flow of class 1 on arc $(i,j)$. Constraints (6.3) imply passengers' demand satisfaction for each $od$ pair, while (6.4) are non-negativity constraints regarding traffic flows. Moreover, constraints (6.5) define the relation among $f_l^{od,1}$ and $x_{i,j}^1$ by means of the path-arc incidence matrix whose elements are defined as follows:

$$\delta_{i,j,l}^{od,1} = \begin{cases} 1 & \text{if } (i,j) \text{ belongs to path } l \text{ from } o \text{ to } d \\ 0 & \text{otherwise} \end{cases} \tag{6.6}$$

Note that (6.3)-(6.5) are equivalent to (2.9)-(2.11) but expressed with respect to the nodes instead of with respect to the links.

In (6.2), the terms $\tau_{i,j}^1(\cdot)$ are the performance functions of arcs related to class $c = 1$, i.e. passengers. These functions, representative of the functional characteristics of the network arcs, express the relation between the travel time spent by passengers traveling through an arc and the amount of congestion on the same arc, which is expected to perform worse (and thus resulting in increased travel times) if the number of users traveling on it increases. The performance functions adopted for this model are derived from (6.36), with $c = 1$, introduced in section 6.4, considering the definitions provided in (6.37) and (6.40). For highway and railway arcs, this equation computes the transit time as a function of the estimated speed. The average speed on these arcs, in turn, depends on the number of users who are using those arcs. Regarding intermodal passenger arcs, on the other hand, the transfer time is considered constant and independent of the number of users present on it. For further information the reader is referred to the description of equations (6.36), (6.37) and (6.40) in section 6.4. It is worth noting that, with the exception of the intermodal case, the link performance functions are strictly increasing hyperbolic functions and therefore diverging in proximity to the arc theoretical maximum capacity.

In order to reduce the computational effort, appropriate linear versions of the same functions have been used within the assignment process. In the case of highway arcs, the resulting linear functions are obtained by interpolating two points: the first is obtained using the free-flow travel time corresponding to an arc completely empty, while the other uses the value assumed by the hyperbolic function when the number of users is equal to $\phi \cdot n_{i,j}^{\max}$, where $\phi \in [0, 1)$. The linear performance functions in the highway case are therefore defined as follows

$$\tau_{i,j}^{1}(x_{i,j}^{1}) = \frac{\Delta_{i,j}}{\omega_{i,j} n_{i,j}^{\max}(1-\phi)} \cdot x_{i,j}^{1} \frac{1}{\eta} + \frac{\Delta_{i,j}}{v_{i,j}^{\mathrm{H}}} \tag{6.7}$$

for $(i, j) \in \mathcal{A}^{\mathrm{H}}$.

Similarly, for the railway case, the second interpolating point is associated with the value assumed by the hyperbolic function when the number of users reaches the technical limit $\frac{C^{\mathrm{P}}\Delta_{i,j}}{s_{i,j}^{\min}}$ (see (6.40)). The following linear performance function for the railway case is therefore obtained:

$$\tau_{i,j}^{1}(x_{i,j}^{1}) = \frac{h_{i,j} s_{i,j}^{\min}}{(s_{i,j}^{\min} - L)} \cdot \frac{x_{i,j}^{1}}{C^{\mathrm{P}}} + \frac{\Delta_{i,j}}{v_{i,j}^{\mathrm{R}}} \tag{6.8}$$

for $(i, j) \in \mathcal{A}^{\mathrm{R}}$.

Finally, as mentioned above, the performance functions of intermodal arcs are constant functions, but it has been necessary to make them strictly increasing by introducing a (although very small) relation of direct proportionality between travel time and number of users on the arc in order to guarantee the uniqueness of the solution of the optimization problem, as detailed further on.

The performance functions for intermodal arcs are then defined as follows

$$\tau_{i,j}^{1}(x_{i,j}^{1}) = \alpha_{i,j} \cdot T + \frac{1}{M} x_{i,j}^{1} \tag{6.9}$$

for $(i, j) \in \mathcal{A}^{\mathrm{Ip}}$, where $\alpha_{i,j} \geq 1$ and $M$ is a positive coefficient sufficiently big such that the impact of the number of users on the performance of an intermodal arc is negligible compared to the other types of arcs.

By defining the performance functions of the arcs as above, it can be proven that Problem 1 is strictly convex on a convex domain and therefore admits a unique optimal solution with respect to variables $x^1_{i,j}$ but not for variables $f_l^{od,1}$. Since $f_l^{od,1}$, properly converted into splitting rates, represent the input of the dynamic model, their uniqueness is an essential requirement in this work. Several methodologies have been developed to overcome this issue associated with the User-Equilibrium model (Borchers et al., 2015). A possible solution is looking for a pattern of flows **f**, coherent with the distribution of users at the equilibrium, maximizing an entropy function and for this reason more likely to occur (Rossi, McNeil, and Hendrickson, 1989).

The resulting optimization problem is as follows.

**Problem 6.**

$$min\ h(\boldsymbol{f}) = \sum_{o \in J^O} \sum_{d \in J^D} \sum_{l \in \mathcal{P}^{od}} f_l^{od,1} \cdot \ln(f_l^{od,1}) \tag{6.10}$$

*subject to*

$$\sum_{l \in \mathcal{P}^{od}} f_l^{od,1} = D^{od,1} \qquad o \in J^O,\ d \in J^D \tag{6.11}$$

$$x_{i,j}^{1,\mathbf{UE}} = \sum_{o \in J^O} \sum_{d \in J^D} \sum_{l \in \mathcal{P}^{od}} f_l^{od,1} \cdot \delta_{i,j,l}^{od,1} \qquad (i,j) \in \mathcal{A} \tag{6.12}$$

Equations (6.11) – (6.12) convey the same constraints of Problem 1 with the only but fundamental difference that the flows on the arcs are now fixed and equal to $x_{i,j}^{1,\mathbf{UE}}$, obtained as solution of Problem 1. Also, the non-negativity constraint (6.4) is now implicit in the fact that $f_l^{od,1}$ appear in (7.39) as the argument of a logarithm.

### 6.3.2 The multi-modal assignment model for freight flows

In this section, the multi-modal assignment model for freight flows is presented. The purpose of this assignment model is to replicate the average freight route choices by considering as objective the minimization of the total travel costs needed to satisfy a given demand. Given these premises, the multi-modal assignment model for freight is given as follows

**Problem 7.**

$$min\ y(\boldsymbol{x}) = \sum_{(i,j)\in\mathcal{A}} x^2_{i,j} \cdot c_{i,j}(x^2_{i,j}) \tag{6.13}$$

*subject to*

$$\sum_{l\in\mathcal{P}^{od}} f^{od,2}_l = D_{od,2} \qquad o \in J^O, d \in J^D \tag{6.14}$$

$$f^{od,2}_l \geq 0 \qquad o \in J^O, d \in J^D, l \in P^{od} \tag{6.15}$$

$$x^2_{i,j} = \sum_{o\in J^O} \sum_{d\in J^D} \sum_{l\in\mathcal{P}^{od,2}} f^{od,2}_l \cdot \delta^{od,2}_{i,j,l} \qquad (i,j) \in \mathcal{A} \tag{6.16}$$

where $f^{od,2}_l$ and $x^2_{i,j}$ are, respectively, the flow of class $c = 2$ of $od$ pair using path $l$ and the total flow of freight on arc $(i,j)$. Constraints (6.14)-(6.16) are analogous to constraints (6.3)-(6.5) included in Problem 1. As done for the passenger assignment, the performance functions that estimate the level of congestion and the travel time on arcs are defined by linearizing (6.36), with (6.37) and (6.40), where $c = 2$. Therefore, for highway arcs, the performance function for freight is defined as follows

$$\tau^2_{i,j}(x^2_{i,j}) = \frac{\Delta_{i,j}}{\omega_{i,j} n^{\max}_{i,j}(1-\phi)} \cdot \varsigma x^2_{i,j} + \bar{t}^1_{i,j} \tag{6.17}$$

where $\varsigma$ is the conversion factor adopted to express the number of trucks $x^2_{i,j}$ in the arcs $(i,j) \in \mathcal{A}^{\mathrm{H}}$ in terms of passenger car equivalents. Equation (6.17) estimates the marginal contribution on travel times due to the presence of freight vehicles, which is summed to $\bar{t}^1_{i,j}$, that is the average travel time over arc $(i,j)$ due exclusively to the presence of passengers, as calculated by the dynamic model. The term $\frac{\Delta_{i,j}}{v^{\mathrm{H}}_{i,j}}$ does not appear explicitly, since the free-flow travel time is contained in $\bar{t}^1_{i,j}$. In fact, it is assumed that the maximum speed $v^{\mathrm{H}}_{i,j}$ is the same for both freight and passenger vehicles.

In the case of railway arcs, a similar approach is adopted and the relative performance function is defined as described below

$$\tau^2_{i,j}(x^2_{i,j}) = \frac{h_{i,j}s^{\min}_{i,j}}{(s^{\min}_{i,j} - L)} \cdot \frac{x^2_{i,j}}{C^{\mathrm{f}}} + \bar{t}^1_{i,j} \tag{6.18}$$

which estimates the marginal contribution to the travel time of the arc due to the presence of freight, summed to the travel time due exclusively to the presence of passengers $\bar{t}^1_{i,j}$. The variable $x^2_{i,j}$ expresses the amount of cargo units present on the arc. Therefore, it follows that $\frac{x^2_{i,j}}{C^{\mathrm{f}}}$ represents the equivalent number of freight trains. The free-flow travel time $\frac{\Delta_{i,j}}{v^{\mathrm{R}}_{i,j}}$ is implicitly contained in $\bar{t}^1_{i,j}$.

The performance functions for freight intermodal arcs $(i,j) \in \mathcal{A}^{\mathrm{If}}$ are defined as follows:

$$\tau^2_{i,j}(x^2_{i,j}) = \gamma_{i,j} \cdot T + \frac{1}{M}x^2_{i,j} \tag{6.19}$$

where $\gamma_{i,j} \geq 1$ and $M$ is a positive coefficient large enough so that the impact due to the presence of cargo units on the arc is insignificant. Given the performance functions for the freight transport of each arc, the cost function in (6.13) associated with each arc of the network is defined as

$$c_{i,j}(x^2_{i,j}) = \tau^2_{i,j}(x^2_{i,j})C^{\mathrm{time}}_{i,j} + \Delta_{i,j}C^{\mathrm{space}}_{i,j} + C^{\mathrm{fix}}_{i,j} \tag{6.20}$$

where $C^{\mathrm{time}}_{i,j}$, $C^{\mathrm{space}}_{i,j}$ and $C^{\mathrm{fix}}_{i,j}$ are the cost per time unit, cost per space unit, and the fixed cost of arc $(i,j)$, respectively. Depending on the type of arc, the contribution made by each of the three parameters can change significantly. For example, in the case of an intermodal freight arc, it is reasonable to assume $C^{\mathrm{space}}_{i,j} = 0$.

Being the performance functions (6.17)-(6.19) monotonically increasing with respect to the freight flows and being the cost function (6.20) linear with respect to the travel times, it follows that also the cost function of the arcs is monotonically increasing with respect to the freight flows. This makes the function $y(\cdot)$ strictly convex and defined on a convex set (6.14)-(6.16) admitting a single optimal solution with respect to the variables $x^2_{i,j}$. However, also for the route choices for freight flows, the uniqueness of the solution is not guaranteed with respect to the flows on paths $f^{od,2}_l$.

To overcome this issue, a problem analogous to Problem 2 can be formulated to obtain the pattern of freight flows on the routes most consistent with the routing problem described in Problem 3.

## 6.4 The macroscopic multi-modal transport network model for freight and passenger flows

The dynamic model adopted in this work is used to capture static and dynamic features of the overall system through a set of discrete-time equations. Aggregate discrete-time models have already been used for performance evaluation and optimization of specific multi-modal transportation processes. In Caballini et al. (2013) and Caballini et al. (2016), for instance, discrete-time models for freight movements by rail in maritime terminals are described. The model presented in this section is much more extensive considering the movement not only of freight but also of passengers and considering a more general applicative context. that is a multi-modal multi-class transport network.

The system dynamic evolution is described by means of aggregate variables defined for each class $c = 1, 2$, for each arc $(i, j) \in \mathcal{A}$, for each $od$ pair, with $o \in J^{\mathrm{O}}$, $d \in J^{\mathrm{D}}$, and for each time step $k$, $k = 0, \ldots, K$. The main aggregate variables adopted in the model are listed below:

- $n_{i,j}^{od,c}(k)$ is the number of units of class $c$ in arc $(i, j)$ associated with the $od$ pair at time step $k$;
- $I_{i,j}^{od,c}(k)$ is the number of units of class $c$ entering arc $(i, j)$ associated with the $od$ pair at time step $k$;
- $O_{i,j}^{od,c}(k)$ is the number of units of class $c$ exiting arc $(i, j)$ associated with the $od$ pair at time step $k$;
- $\beta_{i,j}^{od,c}(k)$ are the splitting rates of class $c$ in arc $(i, j)$ associated with the $od$ pair at time step $k$; note that the condition $\sum_{j \in S(i)} \beta_{i,j}^{od,c}(k) = 1$ must be verified $\forall i, \forall o, \forall d, \forall c, \forall k$.

It is worth clarifying that different units are adopted in the model, depending on the arc type and the flow class. Specifically:

- for class $c = 1$, i.e. passengers, the unit considered in railway arcs is the number of passengers, while in highway arcs is the number of vehicles; in intermodal arcs the units can be either passengers or vehicles depending on the type of arc preceding the intermodal one;
- for class $c = 2$, i.e. freight, the unit considered in railway arcs is the number of railway wagons, while in highway arcs is the number of trucks; in intermodal arcs, again, the units depend on the type of preceding arc. In this work a single cargo unit corresponds to one rail wagon and to one truck. Extending this model for considering different load capacities in road and rail modes is straightforward and omitted here only for the sake of simplicity.

As mentioned in Section 6.2, the dynamic model receives as inputs the route choices of passengers and freight, i.e. the splitting rates $\beta_{i,j}^{od,c}(k)$. These splitting rates $\beta_{i,j}^{od,c}(k)$, for each user class $c$, are obtained from the flows $f_l^{od,c}$ resulting from the application of the multi-modal and multi-class assignment procedure i.e.

$$\beta_{i,j}^{od,c}(k) = \frac{\sum_{l \in \mathcal{P}^{od,c}} f_l^{od,c} \cdot \delta_{i,j,l}^{od,c}}{\sum_{p \in P(i)} \sum_{l \in \mathcal{P}^{od,c}} f_l^{od,c} \cdot \delta_{p,i,l}^{od,c}} \tag{6.21}$$

for all $k$, with $k = 0 \ldots, K$. Note that $\beta_{i,j}^{od,c}(k)$ are constant along the simulation horizon since, in this work, $D^{od,c}$ is the total demand of the whole horizon. In case the simulation horizon is divided in different time intervals, each one characterized by a different demand, the multi-modal assignment model described in Section 6.3 is applied for each time interval, resulting in different splitting rates $\beta_{i,j}^{od,c}(k)$.

Virtual queues at the origin nodes are considered in order to model the presence of flows that have to wait before entering the network. At this purpose, for each time step $k$, $k = 0, \ldots, K$, the following variables are introduced:

- $q^{od,c}(k)$ is the number of units of class $c$ associated with the $od$ pair, that can actually enter the network from node $o \in J^{O}$;

- $l^{od,c}(k)$ is the queue length of class $c$, associated with the $od$ pair, which waits at the origin node $o \in J^{\mathrm{O}}$.

The dynamic evolution of the system is described, for each class $c$ and for each $k$, with $k = 0, \ldots, K$, by the following dynamic equation

$$n_{i,j}^{od,c}(k+1) = n_{i,j}^{od,c}(k) + I_{i,j}^{od,c}(k) - O_{i,j}^{od,c}(k) \tag{6.22}$$

for all $(i,j) \in \mathcal{A}$, $o \in J^{\mathrm{O}}$, $d \in J^{\mathrm{D}}$. Let us now describe separately the entering flows $I_{i,j}^{od,c}(k)$ and the exiting flows $O_{i,j}^{od,c}(k)$.

### Entering flows

The entering flows $I_{i,j}^{od,c}$ used in (6.22) are given by

$$I_{i,j}^{od,c}(k) = \beta_{i,j}^{od,c}(k) \left[ \sum_{n \in P(i)} \epsilon_{n,i}^{c} \cdot O_{n,i}^{od,c}(k) + \xi_{i,j}^{c} \cdot q^{od,c}(k) \right] \tag{6.23}$$

meaning that the flows of each user class $c$ that enter a generic arc $(i,j)$ of the network are the flows that actually succeed in exiting from the previous arcs plus the flows that actually manage to enter from node $i$ if this is also an origin node $o$. These flows are then multiplied for $\beta_{i,j}^{od,c}(k)$, which indicates the portion that decides to use arc $(i,j)$ to reach destination $d$.

Since we are describing the behavior of two class of users in a multi-modal transport network, some parameters necessary to quantify the effective traffic load in the network have to be introduced, i.e. the two conversion factors $\epsilon_{n,i}^{c}$ and $\xi_{i,j}^{c}$ used in (6.23). Starting by class $c = 1$, i.e. passengers, and considering that the transition between two different modes of transport can only occur in an intermodal arc, the conversion factor $\epsilon_{n,i}^{1}$ is given by

$$\epsilon_{n,i}^{1} = \begin{cases} 1 & \text{if } (n,i) \in \mathcal{A}^{\mathrm{H}} \cup \mathcal{A}^{\mathrm{R}} \\ \eta & \text{if } (n,i) \in \mathcal{A}^{\mathrm{Ip}} \text{ and } (i,j) \in \mathcal{A}^{\mathrm{R}} \\ \frac{1}{\eta} & \text{if } (n,i) \in \mathcal{A}^{\mathrm{Ip}} \text{ and } (i,j) \in \mathcal{A}^{\mathrm{H}} \end{cases} \tag{6.24}$$

where $\eta$ is the average number of passengers per car, used to translate the number of vehicles in passengers and viceversa.

The conversion factor $\xi^1_{i,j}$ is instead defined considering that the passenger demand is given in terms of number of passengers and that an origin node cannot be followed by an intermodal arc, therefore

$$\xi^1_{i,j} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{A}^{\mathrm{R}} \\ \frac{1}{\eta} & \text{if } (i,j) \in \mathcal{A}^{\mathrm{H}} \end{cases} \tag{6.25}$$

As for class $c = 2$, i.e. the class referring to freight, both conversion factors $\epsilon^2_{n,i}$ and $\xi^2_{i,j}$ are set equal to 1 because, as mentioned above, a cargo unit is assumed to correspond to one truck and to one rail wagon.

Note that the flows that actually enter an arc $(i,j)$ depend on the capability of the arc to receive flows, i.e., the residual capacity $q^{\mathrm{res}}_{i,j}(k)$, calculated based on the total number of units in the arc $(i,j)$ at the time step $k$. To this end let us define with $n^{\mathrm{tot}}_{i,j}(k)$ the total number of vehicles (expressed in terms of passenger car equivalents) present in arc $(i,j) \in \mathcal{A}^{\mathrm{H}}$ at time step $k$ and with $N^{\mathrm{tot}}_{i,j}(k)$ the total number of trains in arc $(i,j) \in \mathcal{A}^{\mathrm{R}}$ at time step $k$, which are computed as follows

$$n^{\mathrm{tot}}_{i,j}(k) = \sum_{o \in J^{\mathrm{O}}} \sum_{d \in J^{\mathrm{D}}} n^{od,1}_{i,j}(k) + \sum_{o \in J^{\mathrm{O}}} \sum_{d \in J^{\mathrm{D}}} \varsigma n^{od,2}_{i,j}(k) \tag{6.26}$$

for all $(i,j) \in \mathcal{A}^{\mathrm{H}}$, where $\varsigma$ is a coefficient introduced in order to translate the trucks in an equivalent number of cars, and

$$N^{\mathrm{tot}}_{i,j}(k) = \frac{\sum_{o \in J^{\mathrm{O}}} \sum_{d \in J^{\mathrm{D}}} n^{od,1}_{i,j}(k)}{C^{\mathrm{p}}} + \frac{\sum_{o \in J^{\mathrm{O}}} \sum_{d \in J^{\mathrm{D}}} n^{od,2}_{i,j}(k)}{C^{\mathrm{f}}} \tag{6.27}$$

for all $(i,j) \in \mathcal{A}^{\mathrm{R}}$. The total number of vehicles and the total number of trains present in each arc at each time step must ensure the following conditions: $0 \leq n^{\mathrm{tot}}_{i,j}(k) \leq n^{\max}_{i,j}$ and $0 \leq N^{\mathrm{tot}}_{i,j}(k) \leq N^{\max}_{i,j}$.

The residual capacity of highway and railway arcs $q^{\mathrm{res}}_{i,j}(k)$ is given by

$$q^{\mathrm{res}}_{i,j}(k) = \begin{cases} n^{\max}_{i,j} - n^{\mathrm{tot}}_{i,j}(k) & \text{if } (i,j) \in \mathcal{A}^{\mathrm{H}} \\ N^{\max}_{i,j} - N^{\mathrm{tot}}_{i,j}(k) & \text{if } (i,j) \in \mathcal{A}^{\mathrm{R}} \end{cases} \tag{6.28}$$

As for the intermodal arcs $(i,j) \in \mathcal{A}^{\mathrm{Ip}} \cup \mathcal{A}^{\mathrm{If}}$, being fictitious arcs, we chose to not impose bounds on the capacity.

Now let us analyze the receptive capacity of an arc according to the load conditions it is experiencing in a given time step. Then, let us define with $w_{i,j}^{od,c}(k)$ the amount of users of class $c$ related to the $od$ pair who want to enter arc $(i,j)$ in time step $k$

$$w_{i,j}^{od,c}(k) = \beta_{i,j}^{od,c}(k) \left[ \sum_{n \in P(i)} \epsilon_{n,i}^{c} \cdot S_{n,i}^{od,c}(k) + \xi_{i,j}^{c} \cdot \left( \frac{D^{od,c}}{K} + l^{od,c}(k) \right) \right] \quad (6.29)$$

In (6.29), $w_{i,j}^{od,c}(k)$ includes the potential outflow from the previous arcs $S_{n,i}^{od,c}(k)$, better detailed in Section 6.4, and the demand associated with the pair $od$ and the eventual queue length at the node $i$ if it coincides with the origin $o$.

Therefore, the total amount of units that potentially enter arc $(i,j)$ at time step $k$ is given by:

$$W_{i,j}^{\text{tot}}(k) = \begin{cases} \sum_{o \in J^O} \sum_{d \in J^D} w_{i,j}^{od,1}(k) + \varsigma w_{i,j}^{od,2}(k) & \text{if } (i,j) \in \mathcal{A}^H \\ \sum_{o \in J^O} \sum_{d \in J^D} \frac{w_{i,j}^{od,1}(k)}{C^{\text{p}}} + \frac{w_{i,j}^{od,2}(k)}{C^{\text{f}}} & \text{if } (i,j) \in \mathcal{A}^R \end{cases} \quad (6.30)$$

Hence, thanks to the residual capacity defined in (6.28) we can determine the percentage of excess units $\pi_{i,j}(k)$ that cannot enter at a given time step $k$ in the arc $(i,j)$

$$\pi_{i,j}(k) = \begin{cases} \frac{\max\{0, W_{i,j}^{\text{tot}}(k) - q^{\text{res}}(k)\}}{W_{i,j}^{\text{tot}}(k)} & \text{if } W_{i,j}^{\text{tot}}(k) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (6.31)$$

where $\pi_{i,j}(k) \in [0,1]$. Specifically, $\pi_{i,j}(k)$ is equal to 0 when $W_{i,j}^{\text{tot}} < q^{\text{res}}(k)$, i.e. when the residual capacity is sufficient to host all units that want to enter arc $(i,j)$, while $\pi_{i,j}(k)$ is equal to 1 when $q^{\text{res}}(k) = 0$, i.e. when the residual capacity is zero and all units that want to enter arc $(i,j)$ cannot do so.

Having said that, the units that can effectively enter from an origin node is computed as follow

$$q^{od,c}(k) = \left[\frac{D^{od,c}}{K} + l^{od,c}(k)\right] \sum_{n \in S(o)} \beta_{o,n}^{od,c}(k) \cdot (1 - \pi_{o,n}(k)) \tag{6.32}$$

whereas the relative queue of units at the origin node $o$ is given by

$$l^{od,c}(k+1) = l^{od,c}(k) + \frac{D^{od,c}}{K} - q^{od,c}(k) \tag{6.33}$$

**Exiting flows**

Even with respect to outflows from an arc, these are calculated based on the residual capacity of the arcs they wish to enter. For this reason, we distinguish $O_{i,j}^{od,c}(k)$, i.e. the units actually exiting arc $(i,j)$, from $S_{i,j}^{od,c}(k)$, i.e. the units that would like to exit arc $(i,j)$ at time step $k$.

Let us start by describing the relation that defines the potential outflow from an intermodal freight arc $(i,j) \in \mathcal{A}^{\mathrm{If}}$ preceded by a highway arc. We assume that cargo units can only enter the rail network if they are enough to fully load at least a freight train with capacity $C^{\mathrm{f}}$. Since the transfer of cargo units from road to rail is realized through a freight intermodal arc, we consider that this arc behaves as a buffer where cargo units wait until their number is sufficient to fill at least one train and then leave the intermodal arc. The number of cargo units possibly leaving a road-to-rail intermodal arc $(i,j) \in \mathcal{A}^{\mathrm{If}}$ is then given by

$$S_{i,j}^{od,2}(k) = \left\lfloor \frac{n_{i,j}^{od,2}(k)}{C^{\mathrm{f}}} \right\rfloor C^{\mathrm{f}} \tag{6.34}$$

Now let us discuss the potential outflow for all classes $c$ for the arcs $(i,j) \in \mathcal{A}^{\mathrm{H}} \cup \mathcal{A}^{\mathrm{R}}$, for passengers in arcs $(i,j) \in \mathcal{A}^{\mathrm{Ip}}$ and for freight in rail-to-road intermodal arcs $(i,j) \in \mathcal{A}^{\mathrm{If}}$. The potential outflow $S_{i,j}^{od,c}(k)$ is computed as

$$S_{i,j}^{od,c}(k) = \frac{T}{t_{i,j}(k)} n_{i,j}^{od,c}(k) \tag{6.35}$$

where $t_{i,j}(k)$ is the transfer time required to cover arc $(i, j)$ defined according to the following relation

$$t_{i,j}(k) = \begin{cases} \frac{\Delta_{i,j}}{V_{i,j}(n^{\text{tot}}_{i,j}(k))} & \text{if } (i,j) \in \mathcal{A}^{\text{H}} \\ \frac{\Delta_{i,j}}{V_{i,j}(N^{\text{tot}}_{i,j}(k))} & \text{if } (i,j) \in \mathcal{A}^{\text{R}} \\ \alpha_{i,j} \cdot T & \text{if } (i,j) \in \mathcal{A}^{\text{Ip}} \\ \gamma_{i,j} \cdot T & \text{if } (i,j) \in \mathcal{A}^{\text{If}} \end{cases} \tag{6.36}$$

For intermodal connections allowing the modal change from rail to road, the transfer time $t_{i,j}(k)$ is considered constant and equal to $\alpha_{i,j} \cdot T$, with $\alpha_{i,j} \geq 1$ for each $(i, j) \in \mathcal{A}^{\text{Ip}}$ and equal to $\gamma_{i,j} \cdot T$, with $\gamma_{i,j} \geq 1$ for each arc $(i, j) \in \mathcal{A}^{\text{If}}$. Instead, in intermodal connections allowing the road-to-rail modal change, the transfer time is not fixed and depends on the possibility to fill trains, i.e. on (6.34).

With regard to highway and railway arcs, it should be noted that, for both types of arc, the transfer time is estimated as a function of the total number of vehicles or trains present in the connection. More in details, for each highway arc $(i, j) \in \mathcal{A}^{\text{H}}$, the transfer time $t_{i,j}(k)$ is computed according to the current traffic conditions through the steady-state relation between speed and number of vehicles given by

$$V_{i,j}(n^{\text{tot}}_{i,j}(k)) = \min \left\{ v^{\text{H}}_{i,j}, \frac{w_{i,j}}{n^{\text{tot}}_{i,j}(k)} \Delta_{i,j} \left[ \frac{n^{\max}_{i,j}}{\Delta_{i,j}} - \frac{n^{\text{tot}}_{i,j}(k)}{\Delta_{i,j}} \right] \right\} \tag{6.37}$$

Relation (6.37) has been derived from a triangular fundamental diagram, as the one proposed by Daganzo (1994), and expressed in terms of number of vehicles.

Starting from a graphical train timetable, a function expressing the steady-state speed with respect to the number of trains has been defined. Before introducing the relation $V_{i,j}(N^{\text{tot}}_{i,j}(k))$ adopted in (6.36), let us briefly sketch the main steps that have led to the definition of this relation. Let us consider a generic graphical train timetable and let us define with $\Delta X$ the spatial interval, $\Delta T$ the time interval and with $N$ the number of trains present in that frame, hence the average space headway $\bar{s}$ can be defined as $\bar{s} = \frac{\Delta X}{N}$ while the average time headway $\bar{h}$ is given by $\bar{h} = \frac{\Delta T}{N}$. Let us also define the relation between the average space headway and the average time headway that is given by $\bar{s} = \bar{h} \cdot v + L$, where $v$ is the train speed and $L$ is the average length of trains.

Inspired by the traffic fundamental diagram, we can assume that, for a given average time headway, the maximum flow of trains in an arc corresponds to an average space headway equal to $\bar{s} = \bar{h} \cdot v^{\max} + L$, where $v^{\max}$ is the maximum speed allowed. At the same time, it has to be remembered that trains have to maintain a minimum distance $s^{\min}$ that allows them to stop safely; consequently, as the number of trains on an arc increases, they have to reduce their speed in order to guarantee this safety distance. Then, a triangular traffic fundamental diagram in the railway context can be defined as follows

$$Q\left(\frac{1}{\bar{s}}\right) = \begin{cases} \frac{1}{\bar{s}} \cdot v^{\max} \text{ if } \frac{1}{\bar{s}} \leq \frac{1}{\bar{h} \cdot v^{\max} + L} \\ \frac{1}{\bar{h}}\left(1 - \frac{L}{\bar{s}}\right) \\ \quad \text{if } \frac{1}{\bar{h} \cdot v^{\max} + L} < \frac{1}{\bar{s}} \leq \frac{1}{s^{\min}} \end{cases} \tag{6.38}$$

where $Q\left(\frac{1}{\bar{s}}\right)$ is the "flow of trains" in function of the average space headway. The corresponding steady-state speed-headway relation $V\left(\frac{1}{\bar{s}}\right)$ is given by

$$V\left(\frac{1}{\bar{s}}\right) = \begin{cases} v^{\max} \quad \text{ if } \frac{1}{\bar{s}} \leq \frac{1}{\bar{h} \cdot v^{\max} + L} \\ \frac{1}{\bar{h}}(\bar{s} - L) \\ \quad \text{if } \frac{1}{\bar{h} \cdot v^{\max} + L} < \frac{1}{\bar{s}} \leq \frac{1}{s^{\min}} \end{cases} \tag{6.39}$$

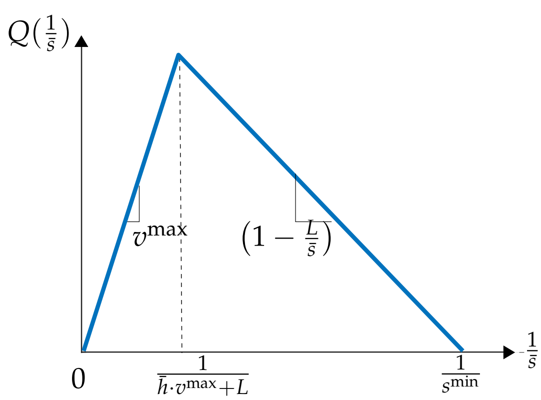



FIGURE 6.3: Triangular fundamental diagram for railway traffic

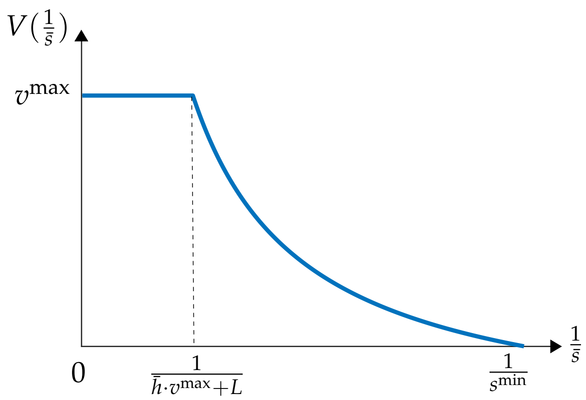



FIGURE 6.4: Steady-state speed-headway functions for railway traffic

By considering that the inverse of the average space headway has an equivalent meaning of "density of trains" in an arc, i.e. $\frac{N_{i,j}^{\text{tot}}(k)}{\Delta_{i,j}}$, the steady-state speed relation $V_{i,j}(N_{i,j}^{\text{tot}}(k))$ for all $(i,j) \in \mathcal{A}^{\text{R}}$ may be formulated as

$$V_{i,j}(N_{i,j}^{\text{tot}}(k)) = \begin{cases} v_{i,j}^{\text{R}} & \text{if } \frac{N_{i,j}^{\text{tot}}(k)}{\Delta_{i,j}} \leq \frac{1}{h_{i,j}v_{i,j}^{\text{R}}+L} \\ \frac{1}{h_{i,j}}\left(\frac{\Delta_{i,j}}{N_{i,j}^{\text{tot}}(k)} - L\right) & \\ & \text{if } \frac{1}{h_{i,j}v_{i,j}^{\text{R}}+L} < \frac{N_{i,j}^{\text{tot}}(k)}{\Delta_{i,j}} \leq \frac{1}{s_{i,j}^{\min}} \end{cases} \tag{6.40}$$

It is worth noting that, for each highway or railway arc, condition (6.1) with (6.37) and (6.40) implies that the transfer time $t_{i,j}(k)$ is never lower than $T$, ensuring the validity of the conservation equations.

Finally, given $S_{i,j}^{od,c}(k)$ and $\pi_{i,j}(k)$, we can compute the outflow $O_{i,j}^{od,c}(k)$ which represent the units of class $c$ referred to the pair $od$ that actually exit the arc $(i,j)$ as

$$O_{i,j}^{od,c}(k) = S_{i,j}^{od,c}(k) \sum_{n\in S(j)} \beta_{j,n}^{od,c}(k)\big(1-\pi_{j,n}(k)\big) \tag{6.41}$$

## 6.5 Simulation results

The focus of this section is to show the potential benefits that can be obtained by adopting the proposed multi-class multi-modal modeling scheme, simulating a perturbation on the network and then analyzing how this event can affect the rest of the network. Specifically, the considered perturbation is the failure of a connection in the multi-modal network. The experimental tests have been performed in two distinct scenarios:

- *Scenario without disruption*: network at the equilibrium before the advent of the disruption;
- *Scenario with disruption*: network once a new equilibrium is reached some time after the initial perturbation.

To evaluate the performance of the multi-modal transport network before and after the disruption, the *Total Travel Time* referred to each arc is calculated as follows:

$$TTT_{i,j} = T \sum_{k=1}^{K} \sum_{c=1}^{2} \sum_{o \in J^{\mathrm{O}}} \sum_{d \in J^{\mathrm{D}}} n_{i,j}^{od,c}(k) \tag{6.42}$$

The total travel time is an indicator that computes the total time spent by each unit in a connection, considering that an arc can belong to multiple paths at the same time.

The results have been obtained by adopting a test network derived from the well-known Nguyen-Dupuis network properly modified to consider the multi-modal case (for further information see Nguyen and Dupuis (1984)). This network is composed of 14 nodes and 21 arcs, as depicted in Fig. 6.5. The critical event is simulated considering the total loss of functionality of the railway arc 12-8.

The main parameters of the highway and railway arcs are reported in Table 6.6 and Table 6.7, respectively. The other parameters have been set as follows: the conversion factor $\eta$, representing the average number of passengers per car, is equal to 1.45, the congestion wave speed $\omega_{i,j}$ is equal to 30 [km/h], $\forall (i,j) \in \mathcal{A}^{\mathrm{H}}$, the average time headway $h_{i,j}$ is 15 minutes and the minimum average space headway $s_{i,j}^{\mathrm{min}}$ is equal to 2 [km], $\forall (i,j) \in \mathcal{A}^{\mathrm{R}}$, while the constant parameter $\alpha_{i,j}$ has been set equal to 15, $\forall (i,j) \in \mathcal{A}^{\mathrm{Ip}}$ and $\gamma_{i,j}$ equal to 30, $\forall (i,j) \in \mathcal{A}^{\mathrm{If}}$. The capacity of a freight train $C^{\mathrm{f}}$ is 25 rail wagons, while the capacity of a passenger train $C^{\mathrm{p}}$ is chosen equal to 700 passengers.

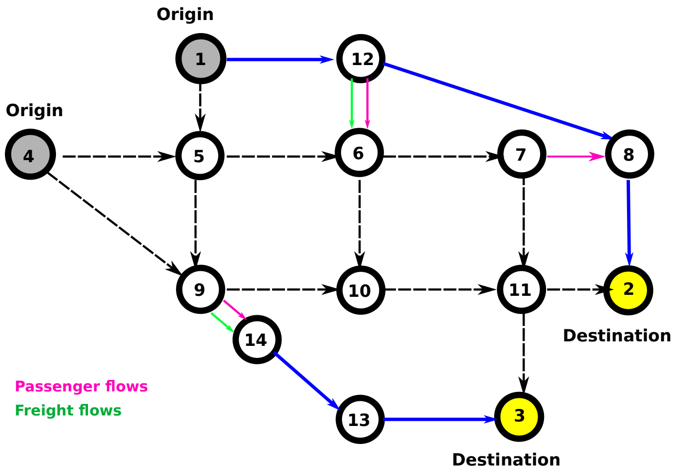



FIGURE 6.5: Sketch of the test multi-modal transport network.

As for the macroscopic multi-modal transport network model, a sample time $T$ equal to one minute has been adopted, while the time horizon of the simulation has been set equal to one hour, that corresponds to $K = 60$ time steps. The transportation demand is expressed with two origin-destination matrices, one referring to passengers, as shown in Table 6.8, and one referring to freight, as indicated in Table 6.9. The mobility demand of both passengers and freight is assumed to be perfectly inelastic and therefore it is the same in the two tested scenarios.

Note finally that, in both scenarios described below, the splitting rates $\beta_{i,j}^{od,c}(k)$ computed by the multi-modal traffic assignment are constant throughout the simulation period, and, then, $k$ is omitted for the sake of simplicity, i.e. $\beta_{i,j}^{od,c}(k) = \beta_{i,j}^{od,c}$.

### 6.5.1 Scenario without disruption

The methodology presented in section 6.3 has been used to allocate the passengers demand on possible routes and the resulting assignment is shown in Fig. 6.6. It is worth reminding that the feasible paths are those that have at most one modal shift. The paths used at the equilibrium, before the disruption, are reported in Table 6.10: one path adopts only the rail mode,

TABLE 6.6: Main parameters referred to highway arcs.

| | $\Delta_{i,j}$ [km] | $v^{\text{H}}_{i,j}$ [km/h] |
|---|---|---|
| **arc 1-5** | 6 | 70 |
| **arc 4-5** | 5 | 70 |
| **arc 4-9** | 12 | 70 |
| **arc 5-6** | 40 | 120 |
| **arc 5-9** | 6 | 70 |
| **arc 6-7** | 40 | 120 |
| **arc 6-10** | 8 | 70 |
| **arc 7-11** | 7 | 70 |
| **arc 9-10** | 40 | 120 |
| **arc 10-11** | 40 | 120 |
| **arc 11-2** | 6 | 70 |
| **arc 11-3** | 4 | 70 |

TABLE 6.7: Main parameters referred to railway arcs.

| | $\Delta_{i,j}$ [km] | $v^{\text{R}}_{i,j}$ [km/h] |
|---|---|---|
| **arc 1-12** | 40 | 100 |
| **arc 8-2** | 6 | 70 |
| **arc 12-8** | 50 | 100 |
| **arc 13-3** | 40 | 70 |
| **arc 14-13** | 45 | 70 |

TABLE 6.8: Passengers origin-destination matrix.

| $D^{od,1}$ | **Destination node 2** | **Destination node 4** |
|---|---|---|
| **Origin node 1** | 1800 | 1500 |
| **Origin node 4** | 2500 | 2000 |

TABLE 6.9: Freight origin-destination matrix.

| $D^{od,2}$ | **Destination node 2** | **Destination node 4** |
|---|---|---|
| **Origin node 1** | 150 | 80 |
| **Origin node 4** | 40 | 25 |

five are highway routes, while the remaining eight require the use of both modes of transport.

TABLE 6.10: Paths used by passengers in the pre-disruption scenario

| *od* **pair** | **paths** |
|---|---|
| **1-2** | [1 12 8 2] |
| **1-3** | [1 5 9 10 11 3], [1 5 6 7 11 3], [1 12 6 7 11 3], [1 5 9 14 13 3] |
| **4-2** | [4 9 10 11 2], [4 5 9 10 11 2], [4 5 6 10 11 2], [4 5 6 7 11 2], |
| **4-3** | [4 9 10 11 3], [4 5 6 7 11 3], [4 9 14 13 3] |

TABLE 6.11: Paths used by freight in the pre-disruption scenario

| *od* **pair** | **paths** |
|---|---|
| **1-2** | [1 12 8 2] |
| **1-3** | [1 5 9 14 13 3] |
| **4-2** | [4 5 6 7 11 2], [4 5 9 10 11 2], [4 9 10 11 2] |
| **4-3** | [4 9 14 13 3], [4 5 9 14 13 3] |

The non-zero splitting rates are in this case:

- *od* pair 1-2: $\beta^{12,1}_{1,12} = 1$, $\beta^{12,1}_{5,6} = 1$, $\beta^{12,1}_{6,7} = 1$,$\beta^{12,1}_{7,11} = 1$,$\beta^{12,1}_{8,2} = 1$, $\beta^{12,1}_{11,2} = 1$, $\beta^{12,1}_{12,8} = 1$;
- *od* pair 1-3: $\beta^{13,1}_{1,5} = 0.92$, $\beta^{13,1}_{1,12} = 0.92$, $\beta^{13,1}_{5,6} = 0.36$, $\beta^{13,1}_{5,9} = 0.64$, $\beta^{13,1}_{6,7} = 1$, $\beta^{13,1}_{7,11} = 1$, $\beta^{13,1}_{9,10} = 0.12$,$\beta^{13,1}_{9,14} = 1$, $\beta^{13,1}_{10,11} = 1$, $\beta^{13,1}_{11,3} = 1$, $\beta^{13,1}_{12,6} = 1$, $\beta^{13,1}_{13,3} = 1$, $\beta^{13,1}_{14,3} = 1$;
- *od* pair 4-2: $\beta^{42,1}_{4,5} = 0.62$, $\beta^{42,1}_{4,9} = 0.38$, $\beta^{42,1}_{5,6} = 0.75$, $\beta^{42,1}_{5,9} = 0.25$, $\beta^{42,1}_{6,7} = 0.67$, $\beta^{42,1}_{6,10} = 0.33$, $\beta^{42,1}_{7,11} = 1$, $\beta^{42,1}_{9,10} = 1$, $\beta^{42,1}_{10,11} = 1$, $\beta^{42,1}_{11,2} = 1$;
- *od* pair 4-3: $\beta^{43,1}_{4,5} = 0.18$, $\beta^{43,1}_{4,9} = 0.82$, $\beta^{43,1}_{5,6} = 1$, $\beta^{43,1}_{6,7} = 1$, $\beta^{43,1}_{7,11} = 1$, $\beta^{43,1}_{9,10} = 0.28$, $\beta^{43,1}_{9,14} = 0.72$, $\beta^{43,1}_{10,11} = 1$, $\beta^{43,1}_{11,3} = 1$, $\beta^{43,1}_{13,3} = 1$, $\beta^{43,1}_{14,13} = 1$.

The freight flows are fixed and their distribution is shown again in Fig. 6.6, while the paths are reported in Table 6.11. The corresponding non-zero splitting rates, again considered constant throughout the simulation, are the following:

- *od* pair 1-2: $\beta^{12,2}_{1,12} = 1$, $\beta^{12,2}_{8,2} = 1$, $\beta^{12,2}_{12,8} = 1$ ;
- *od* pair 1-3: $\beta^{13,2}_{1,5} = 1$, $\beta^{13,2}_{5,9} = 1$, $\beta^{13,2}_{9,14} = 1$, $\beta^{13,2}_{13,3} = 1$, $\beta^{13,2}_{14,13} = 1$;
- *od* pair 4-2: $\beta^{42,2}_{4,5} = 0.75$, $\beta^{42,2}_{4,9} = 0.24$ $\beta^{42,2}_{5,6} = 0.17$, $\beta^{42,2}_{5,9} = 0.82$ $\beta^{42,2}_{6,7} = 1$, $\beta^{42,2}_{7,11} = 1$, $\beta^{42,2}_{11,2} = 1$, $\beta^{42,2}_{9,10} = 1$, $\beta^{42,2}_{10,11} = 1$;
- *od* pair 4-3: $\beta^{43,2}_{4,9} = 0.21$, $\beta^{43,2}_{4,5} = 0.78$, $\beta^{43,2}_{5,9} = 1$, $\beta^{43,2}_{9,14} = 1$, $\beta^{43,2}_{13,3} = 1$, $\beta^{43,2}_{14,13} = 1$.

It should be noted that among these routes, only one is of intermodal type, with a change from road to rail, one route is entirely by railway, while the remaining ones involve only the use of highway arcs.

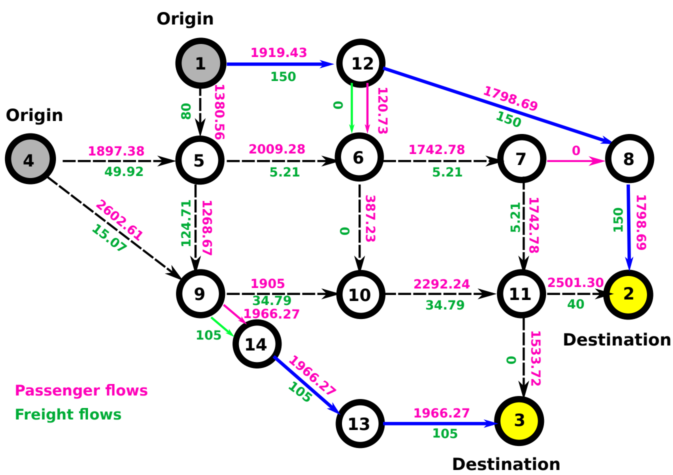



FIGURE 6.6: Pre-disruption equilibrium

### 6.5.2 Scenario with disruption

As mentioned earlier, the disruption is represented in this example by removing the railway arc 12-8. Regarding passengers, the new path configuration for each origin-destination pair is shown in Table 6.12. As can be seen, the railway path [1-12-8-2] no longer appears since it includes the damaged arc. As shown in Fig. 6.7, there is a shift of *od* pair 1-2 flows and, as a consequence, a greater load on the central arcs of the network, in particular on path [1-12-6-7-8-2]. On the other hand, not surprisingly, the southern paths in the network for pair 4-3 are the least perturbed as testified by the fact that the flows do not undergo excessive redistribution, meaning that for those users the most attractive routes have remained the same. The splitting rates for passenger flows are the following:

- *od* pair 1-2: $\beta^{12,1}_{1,5} = 0.41$, $\beta^{12,1}_{1,12} = 0.59$, $\beta^{12,1}_{5,6} = 0.85$, $\beta^{12,1}_{5,9} = 0.15$, $\beta^{12,1}_{6,7} = 0.69$, $\beta^{12,1}_{6,10} = 0.31$, $\beta^{12,1}_{7,8} = 0.27$, $\beta^{12,1}_{7,11} = 0.73$, $\beta^{12,1}_{8,2} = 1$, $\beta^{12,1}_{9,10} = 1$, $\beta^{12,1}_{10,11} = 1$, $\beta^{12,1}_{11,2} = 1$, $\beta^{12,1}_{12,6} = 1$ ;
- *od* pair 1-3: $\beta^{13,1}_{1,5} = 0.52$, $\beta^{13,1}_{1,12} = 0.48$, $\beta^{13,1}_{5,6} = 0.1$, $\beta^{13,1}_{5,9} = 0.9$, $\beta^{13,1}_{6,7} = 0.65$, $\beta^{13,1}_{6,10} = 0.35$, $\beta^{13,1}_{7,11} = 1$, $\beta^{13,1}_{9,14} = 1$, $\beta^{13,1}_{10,11} = 1$, $\beta^{13,1}_{11,3} = 1$, $\beta^{13,1}_{12,6} = 1$, $\beta^{13,1}_{13,3} = 1$, $\beta^{13,1}_{14,3} = 1$;
- *od* pair 4-2: $\beta^{42,1}_{4,5} = 0.57$, $\beta^{42,1}_{4,9} = 0.43$, $\beta^{42,1}_{5,6} = 0.8$, $\beta^{42,1}_{5,9} = 0.2$, $\beta^{42,1}_{6,7} = 0.78$, $\beta^{42,1}_{6,10} = 0.22$, $\beta^{42,1}_{7,8} = 0.54$, $\beta^{42,1}_{7,11} = 0.46$, $\beta^{42,1}_{8,2} = 1$, $\beta^{42,1}_{9,10} = 1$, $\beta^{42,1}_{10,11} = 1$, $\beta^{42,1}_{11,2} = 1$ ;
- *od* pair 4-3: $\beta^{43,1}_{4,5} = 0.19$, $\beta^{43,1}_{4,9} = 0.81$, $\beta^{43,1}_{5,9} = 1$, $\beta^{43,1}_{6,10} = 1$, $\beta^{43,1}_{9,10} = 0.22$, $\beta^{43,1}_{9,14} = 0.78$, $\beta^{43,1}_{10,11} = 1$, $\beta^{43,1}_{11,3} = 1$, $\beta^{43,1}_{13,3} = 1$, $\beta^{43,1}_{14,13} = 1$.

TABLE 6.12: Paths used by passengers after the disruption

| *od* **pair** | **paths** |
|---|---|
| **1-2** | [1 5 9 10 11 2], [1 5 6 7 8 2], [1 12 6 10 11 2], [1 12 6 7 11 2], [1 5 6 10 11 2], [1 5 6 10 11 2] |
| **1-3** | [1 12 6 10 11 3], [1 12 6 7 11 3], [1 5 9 14 13 3] [1 5 6 7 11 3] |
| **4-2** | [4 9 10 11 2], [4 5 6 10 11 2], [4 5 6 7 8 2] [4 5 9 10 11 2], [4 5 6 7 11 2] |
| **4-3** | [4 9 10 11 3], [4 9 14 13 3], [4 5 9 14 13 3] |

As far as freight flows are concerned, similar considerations can be made, since the railway path can no longer be used and the freight flow of *od* pair 1-2 is reassigned to the remaining path, as reported in Table 6.13 with the resulting assignment shown in Fig. 6.7. The corresponding non-zero splitting rates are as follows:

TABLE 6.13: Paths used by freight after the disruption

| *od* pair | paths |
|---|---|
| **1-2** | [1 12 6 10 11 2], [1 12 6 7 11 2] |
| **1-3** | [1 5 9 14 13 3] |
| **4-2** | [4 9 10 11 2], [4 5 9 10 11 2] |
| **4-3** | [4 9 14 13 3], [4 5 9 14 13 3] |

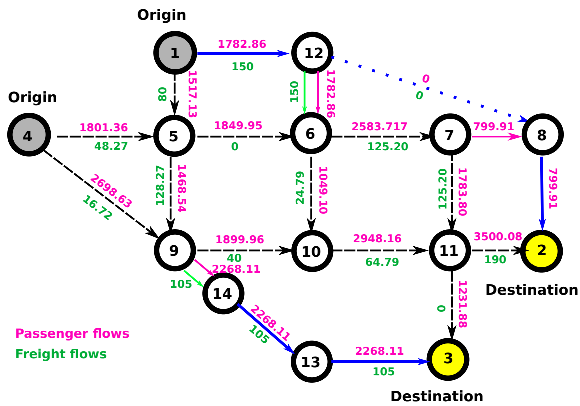



FIGURE 6.7: Post-disruption equilibrium

- *od* pair 1-2: $\beta^{12,2}_{1,12} = 1$, $\beta^{12,2}_{6,7} = 0.83$, $\beta^{12,2}_{6,10} = 0.16$, $\beta^{12,2}_{6,7} = 1$, $\beta^{12,2}_{10,11} = 1$, $\beta^{12,2}_{11,2} = 1$, $\beta^{12,2}_{12,6} = 1$;
- *od* pair 1-3: $\beta^{13,2}_{1,5} = 1$, $\beta^{13,2}_{5,9} = 1$, $\beta^{13,2}_{9,14} = 1$, $\beta^{13,2}_{13,3} = 1$, $\beta^{13,2}_{14,13} = 1$;
- *od* pair 4-2: $\beta^{42,2}_{4,5} = 0.65$, $\beta^{42,2}_{4,9} = 0.35$, $\beta^{42,2}_{5,9} = 1$, $\beta^{42,2}_{6,7} = 1$,$\beta^{42,2}_{7,11} = 1$ $\beta^{42,2}_{9,10} = 1$, $\beta^{42,2}_{10,11} = 1$, $\beta^{42,2}_{11,2} = 1$;
- *od* pair 4-3: $\beta^{43,2}_{4,5} = 0.89$, $\beta^{43,2}_{4,9} = 0.11$, $\beta^{43,2}_{5,9} = 1$, $\beta^{43,2}_{9,14} = 1$, $\beta^{43,2}_{13,3} = 1$, $\beta^{43,2}_{14,13} = 1$.

Table 6.14 shows, for each arc, $TTT^E_{i,j}$, i.e. the total travel time computed in the pre-disruption scenario, $TTT^D_{i,j}$, i.e. the total travel time in the post-disruption scenario, and $\Delta TTT_{i,j}$, i.e. the absolute deviations of the same metric. Analogously, Fig. 6.8 shows, for each arc, the total travel time percent variation by exhibiting the effects of disruption more clearly.

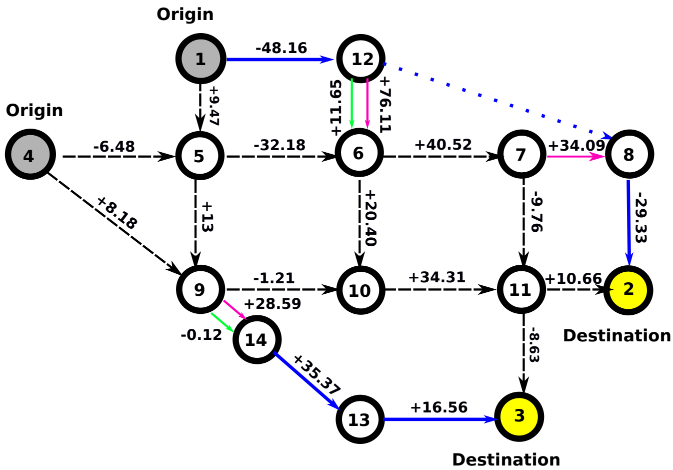



FIGURE 6.8: Total Travel Time absolute variations.

TABLE 6.14: Total travel time for each arc

| | $TTT^E_{i,j}$ | $TTT^D_{i,j}$ | $\Delta TTT_{i,j}$ |
|---|---|---|---|
| **arc 1-5** | 101.33 | 110.80 | +9.47 |
| **arc 1-12** | 629.54 | 581.38 | -48.16 |
| **arc 4-5** | 129.16 | 122.68 | -6.48 |
| **arc 4-9** | 219.33 | 227.51 | +8.18 |
| **arc 5-6** | 392.94 | 360.76 | -32.18 |
| **arc 5-9** | 88.99 | 101.99 | +13.00 |
| **arc 6-7** | 227.88 | 268.40 | +40.52 |
| **arc 6-10** | 20.57 | 40.97 | +20.40 |
| **arc 7-8** | 0.00 | 34.09 | +34.09 |
| **arc 7-11** | 40.92 | 31.17 | -9.76 |
| **arc 8-2** | 38.46 | 9.13 | -29.33 |
| **arc 9-10** | 360.60 | 359.39 | -1.21 |
| **arc 10-11** | 251.67 | 285.98 | +34.31 |
| **arc 11-2** | 52.57 | 63.23 | +10.66 |
| **arc 11-3** | 22.41 | 13.78 | -8.63 |
| **arc 12-8** | 184.64 | - | - |
| **arc 13-3** | 107.92 | 124.49 | +16.56 |
| **arc 14-13** | 230.44 | 265.81 | +35.37 |
| **arc 9-14 If** | 37.96 | 37.84 | -0.12 |
| **arc 9-14 Ip** | 186.27 | 214.86 | +28.59 |
| **arc 12-6 If** | 0.00 | +11.65 | +11.65 |
| **arc 12-6 Ip** | 5.41 | 81.52 | +76.11 |

As can be seen, the disruption of the 12-8 arc implies that freight and passenger flows, which previously used the more peripheral rail route, are re-assigned in more internal routes, particularly multi-modal routes. Consistent with an overall load increase on the central arcs of the network, it is possible to observe a shift, though small, of flows of the 4-3 pair in favor of the outermost multi-modal route [4-9-14-13-3]. Not surprisingly, the arcs relatively most affected by the perturbation are those in proximity to the disrupted arc, such as arcs 12-6, 6-7 and 7-8. However, it can be seen that even the central arc 6-10 experiences a 30% increase in total travel time. The computation of a metric such as the total travel time, by means of the multi-modal transport model presented in this chapter, allows therefore to identify the elements of the network that will be more stressed after the perturbation.

## 6.6 Possible applications of the proposed approach

This section discusses possible applications of the proposed approach. Specifically, the possible uses can be distinguished into *scenario evaluation* and *development and testing of regulatory policies and control actions*. These two application areas are briefly discussed below.

### Scenario evaluation

The proposed model can be adopted to evaluate the effects of the decisions of the users on a transport network. These effects can be quantified through the adoption of specific performance indexes such as the $TTT$ introduced earlier or through the development of other indicators. For instance, this model can also be used to assess the sustainability of user choices by integrating this modeling scheme with models that estimate emissions or energy consumptions.

Scenarios of particular relevance are those concerning the occurrence of critical events. Indeed, large transport networks are anyway susceptible to critical events that may have severe implications on the whole activity system of a territory. These disruptive events may be caused by natural phenomena (such as floods, earthquakes, pandemics, etc.) or anthropogenic causes (such as terrorist attacks, infrastructure failures, but also planned maintenance works): whatever the cause, they may affect the transport network as a variation of mobility demand or as events that partially or totally deteriorate the capacity of a transport network. In this regard, a topic that is attracting particular attention in scientific research is the evaluation of the resilience of a transport network, i.e. its ability to resist, adapt or change in order to maintain acceptable performance in case of critical events. Although the concept of resilience has been initially introduced to describe a property of natural systems (Holling, 1973), recently it has been applied to transport networks and in particular to road networks (Gauthier, Furno, and El Faouzi, 2018; Siri, Siri, and Sacone, 2020a; Siri, Siri, and Sacone, 2020b) and railway networks (Adjetey-Bahun et al., 2016; Dorbritz, 2011; Bešinović, 2020). However, transport networks are complex and highly interdependent systems and a critical event affecting one mode of transport can have an impact on other modes giving rise to a chain effect. What is lacking in the literature, and what this work intends to address, is the possibility of using a tool that allows to quantify the effects of these interdependencies and to evaluate the ability of a transport network to maintain acceptable performance even when it is affected by disruptions.

**Development and testing of regulatory policies and control actions**

The modeling framework proposed in this chapter may constitute the basis for regulation and control approaches finalized at defining routing and modal indications to be provided to the users. First of all, this model can be used to provide more detailed information to users about travel times or route choices for improving sustainability. Moreover, specific routing instructions can be defined for the users and, since the modeling scheme is multi-class, such instructions can be suitably defined for each class of users.

Finally, the proposed two-stage modeling framework may be adopted to test control policies that aim at fully exploiting the mobility capacity of a large-scale multi-modal transport network by suggesting routes that may include one or more transport modes.This can be done both in nominal conditions of the network or when the transport system is affected by an event which changes its structure or the mobility demand.

## 6.7 Conclusions

A two-stage modeling approach is presented to represent passenger and freight flows on a multi-modal transportation network, i.e. a network in which there are roadways, railways, and connections that allow the modal shifts. The modeling scheme consists of an assignment model and a discrete-time macroscopic dynamic model. The assignment model allows to represent the choices of the users, both passengers and freight, in terms of routes and transport modes. These route choices provide the input to the dynamic model that allows to represent the evolution in time and space of user flows, allowing to evaluate some dynamic characteristics such as speed and travel time on the network arcs.

The methodology presented has been tested on a case study in which one of the possible applications of this modeling scheme has been shown. The objective of this analysis has been to evaluate the behavior of a network subject to a disruptive event, specifically the loss of functionality of a railway arc. This application revealed the ability of the multi-modal model to capture the ripple effects of such events that cannot be gained if analyzed with models representing a single transportation mode only.

# Chapter 7

# Bounded rationality-based day-to-day traffic assignment model with topological proximity dependent costs

Based on *Siri, Siri, and Sacone* (*2022a*)

In this chapter, the insights concerning the inertia associated with users' decision making process as introduced in chapter 5 are extended and further explored in a path-based day-to-day traffic assignment model that takes into account users bounded rationality. More specifically, the assignment model belongs to the class of proportional-switch adjustment processes. It is inspired by the discrete implementation of Smith's model (Smith, 1984) formulated by Guo and Huang (2016) and it is specifically designed to represent the evolution of a network under substantial alterations. For any given day, the amount of net flow that shifts from one path to another in response to changing network conditions is computed. As discussed in section 4.1, a common way to interpret the concept of bounded rationality in the context of traffic assignments is by defining an indifference band. A driver is therefore stimulated to change route only when the travel time exceeds a certain threshold. In addition, Lotan (1997) shows that the drivers familiar with the network are less likely to abandon roads they already use. One interpretation of this phenomenon may be found in the empirical study conducted by Hiraoka et al. (2002), where it is shown that drivers tend to prefer routes that are cognitively easier to determine. Choosing to adopt solutions that have already been previously drafted is likely to be cognitively cheaper compared to developing new ones. The empirical study by Vreeswijk et al. (2013) investigates users' route choice behaviour and in particular the underlying reasons for their systematic travel time miscalculation. It is shown that only 41% of respondents choose the actual shortest route. The results of the interviews support the hypothesis according to which users are affected by "choice supportive bias" (Mather, Shafir, and Johnson, 2003), i.e. users are inclined to associate positive attributes to the choices they have made while, conversely, they are more likely to associate negative attributes to the alternatives they have not chosen. Regarding route choice, a clear correlation is shown between travel time overestimation and whether a route was chosen or not. Travel time overestimation tends to be on average significantly higher for non-chosen routes. Meanwhile, it is well recognized in the literature that the extent of overlap between paths can affect their probability of being chosen. For instance, one of the greatest flaws attributed to logit-based stochastic models is that they present two paths as totally distinct alternatives even when they almost completely overlap (independence of irrelevant alternatives). Therefore it is reasonable to assume that the more two paths overlap, the more indifferent a traveller is in using one rather than the other. The model proposed in this chapter, which represents an extension of Siri, Siri, and Sacone (2022b), attempts to encapsulate this behavior with a dynamics in which the users reconsider their travel choices not only based on the congestion level on the routes they are experiencing but also on the degree of topological similarity between the roads potentially improving their condition and the one currently used. When pushed by new network conditions to change routes, users will favor those that more strongly overlap with the one they currently use. In other words, users on different routes perceive the cost of the same route differently. We call this phenomenon as "spatial inertia". This tendency is formally expressed by an over-cost that users assign to the paths they are not using. This approach is similar to the one applied by Cantarella and Cascetta (1995) where, within a conditional path choice model, an extra utility is attributed to the path chosen on the previous day, representing a transition cost. In the present work, however, an extra disutility (extra travel cost) is attributed to the non-chosen routes which is not considered fixed but instead inversely proportional to the topological similarity between a route and the one chosen on the previous day. Additionally, a "temporal inertia" is also introduced to capture the myopia phenomenon identified by Conlisk (1996). The users with a travel cost substantially decreased as a result of a switch will be less inclined to eventually leave it in the future since they are confident of the improvements already achieved. We will show how these attitudes correspond to a bounded rationality behavior and we will prove that the stationary point reached by the dynamic process corresponds to a BRUE.

Recently, Zhang et al. (2019) proposed a continuous-time DTD which takes into account the path overlap within the switch process. More in detail, both the cost difference between the paths moderated by a variable indifference band, which scales with respect to the costs, and the net marginal cost, intended as the marginal cost difference between the paths, are taken into account in the user decision process. According to Zhang et al. (2019), the users prefer, given the same amount of actual travel cost savings, the routes that will result in a lower cost increase once the switch takes place. As a result, given the same travel cost savings, users tend to prefer those routes that overlap with the one currently used, since the overlapped portion does not contribute to the net marginal cost increase. In the present work, conversely, such a behavior is obtained from the topological characteristics of the network exclusively and not from the functional features of the arcs. The flow switch process is not directly influenced by marginal costs, although it is indirectly influenced once the costs are updated (this is a common feature of DTD models), but it is instead clearly influenced by the overlapping percentage of the paths, evaluated taking into account the specific link lengths since links of different lengths have a different impact.

The main features of the proposed proportional-switch day-to-day discrete-time adjustment model can be summarized as follows:

- spatial inertia, i.e. topological proximity dependent costs are introduced to represent that users have a high inertia to change route if the new route has little overlap with the one currently used;
- temporal inertia, i.e. users who have just achieved a good improvement with a switch present a high inertia to change route again;
- the model is based on the interpretation of users' choices and then it is proven that this behavior corresponds to a bounded rational attitude, i.e the stationary point reached by the system is a BRUE.

From a theoretical point of view, the present model proposes a different approach for representing users' bounded rationality. Relying on a unique indifference band, fixed or relative, enables to represent the inertia involved in abandoning an option but it does not provide any insight about the option adopted instead. Answering this question generally results in picking the route that provides the user with the largest travel time/cost savings. This approach implicitly assumes that the user considers each alternative distinct and independent. But the works by Lotan

(1997) and Vreeswijk et al. (2013) suggest how this may not be true. It is therefore reasonable to assume that the inertia involved in abandoning one solution in favor of another should be influenced by the perceived similarity between the alternatives. In other words, small cost/time variations may trigger modest flow adjustments, i.e. between similar paths, while larger cost/time fluctuations may cause larger flow adjustments, leading to a switch even towards paths quite different from those previously used. By adopting this approach, it is possible to represent multiple indifference thresholds, whose size varies according to the extent of correlations between the alternatives (overlapping).

From an applicative point of view, the proposed model can be utilized for the representation of the dynamics emerging out of the interaction between users and the transport infrastructure, especially when the network undergoes significant planned or unintentional alterations. Therefore, by defining appropriate performance metrics, it is possible to assess, for example, the degree of resilience or vulnerability of a network relying not only on topological analysis, but also properly taking into account the impact of user reaction over time.

## 7.1 The Proposed Model

The transport network is represented by an oriented fully-connected graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$ with a finite set $\mathcal{N}$ of nodes and a finite set $\mathcal{A}$ of links. Let $\mathcal{R} \subseteq \mathcal{N}$ and $\mathcal{S} \subseteq \mathcal{N}$ be the subset of origin nodes and the subset of destination nodes, respectively. Time is discretized and represented by the variable $t = t_0, t_1, ...$, indicating a generic day. Note that disruptions cannot occur at $t_0$ but only on subsequent days. This allows to define a reference of the system fundamental quantities in a pre-disruption scenario, required in the definition of appropriate performance metrics, as will be illustrated later. The main notation of the proposed model is reported in Table 7.1.

Let $\mathcal{H} = \{h : h \in \mathcal{R} \times \mathcal{S}\}$ be the set of all origin-destination node pairs while $\mathbf{q} = \{q_h : h \in \mathcal{H}\}$ is the associated travel demand vector, where an element $q_h$ represent the demand for OD pair $h$. Let $\mathcal{K}^t$ denotes the

| Sets | |
|---|---|
| $\mathcal{N}$ | Set of nodes |
| $\mathcal{A}$ | Set of links |
| $t_0, t_1, ...$ | Days |
| $\mathcal{H}$ | Set of origin-destination (OD) pairs |
| $\mathcal{K}_h^t$ | Set of all loop-free paths of OD pair $h$ on day $t$ |
| $\Delta$ | Link-path incidence matrix |
| $\delta_{ak}^h$ | Equal to 1 if link $a$ belongs to path $k$ of OD pair $h$ |
| $\Theta$ | Link-path feasible set |
| $\bar{C}^h(t)$ | Average actual travel cost of OD pair $h$ on day $t$ |
| $E^h(t)$ | Historic average travel cost of OD pair $h$ on day $t$ |

| Flow Vectors | |
|---|---|
| $\mathbf{q}$ | Travel demand vector |
| $q_h$ | Travel demand of OD pair $h \in \mathcal{H}$ |
| $\mathbf{x}(t)$ | Link-flow vector on day $t$ |
| $x_a(t)$ | Flow on link $a \in \mathcal{A}$ on day $t$ |
| $\mathbf{f}(t)$ | Path-flow vector on day $t$ |
| $f_k^h(t)$ | Flow on path $k \in \mathcal{K}_h^t$ on day $t$ |

| Costs | |
|---|---|
| $\mathbf{c}(\mathbf{x}(t))$ | Link cost vector on day $t$ |
| $c_a(x_a(t))$ | Travel cost on path $a \in \mathcal{A}$ on day $t$ |
| $C^h(t)$ | Relative path cost matrix for OD pair $h$ on day $t$ |
| $C_{ks}^h(t)$ | Relative cost of path $s$ compared with path $k$ of OD pair $h$ on day $t$ |
| $S_{ks}^h(t)$ | Cost of switching from path $k$ to path $s$ of OD pair $h$ on day $t$ |
| $A_s^h(t)$ | Actual travel cost on path $s$ of OD pair $h$ on day $t$ |
| $D_{sk}^h(t)$ | Cost-based swap-rate between path $k$ and path $s$ of OD pair $h$ at day $t$ |
| $L^h(t)$ | Myopia-based swap rate of OD pair $h$ on day $t$ |

| Parameters | |
|---|---|
| $\Psi$ | Switch coefficient |
| $\phi$ | Myopia coefficient |
| $\psi$ | Path usage threshold |
| $M$ | Reluctance coefficient |

TABLE 7.1: Main notation

set of all available path connecting each origin with the associated destination on day t. Regarding the flows traveling on the network, the vector $\mathbf{x}(t) = \{x_a(t), a \in \mathcal{A}\}$ represents the link flows, where $x_a(t)$ is the amount of traffic on link $a \in \mathcal{A}$ on day $t$. On the other hand, the vector of path flows is denoted as $\mathbf{f}(t) = \{f_k(t), h \in \mathcal{H}, k \in \mathcal{K}^t\}$, where $f_k(t)$ identifies the traffic flow on path $k \in \mathcal{K}^t$ at day $t$. In order for the flows to be consistent with each other, the following matrices are defined. Let $\Delta = \{\delta_{ak} : a \in \mathcal{A}, k \in \mathcal{K}^t\}$ be the link-path incidence matrix where each element $\delta_{ak} = 1$ if link $a$ is part of path $k$. Finally, let $\Xi = \{\xi_{hk} : h \in \mathcal{K}^t, h \in \mathcal{H}\}$ be the OD-path incidence matrix, where $\xi_{hk} = 1$, if path $k$ connects OD pair $h$. Consequently, for each day, the following relationships must be satisfied

$$\mathbf{x}(t) = \Delta \mathbf{f}(t) \tag{7.1}$$

$$\mathbf{q} = \Xi \mathbf{f}(t) \tag{7.2}$$

$$\mathbf{f}(t) \geq 0 \tag{7.3}$$

where (7.1) requires link-flow and path-flow vectors to be consistent, (7.2) requires demand-flows and path-flows to be consistent while (7.3) implies flows non negativity on each day. Thus the set $\Theta$ of feasible link and path flows can be defined as:

$$\Theta = \{(\mathbf{x}, \mathbf{f}) : (7.6) - (7.8) \text{ hold}\} \tag{7.4}$$

In order to make some of the formulations that will be presented in the following clearer and easier to read, let us to introduce this additional notation. Let $\mathcal{K}_h^t \subseteq \mathcal{K}^t$ be the set of path connecting OD pair $h \in \mathcal{H}$. Consequently, we say that a path belongs to the OD pair $h$, if the following hold.

$$f_k^h(t) = f_k(t) : \xi_{hk} = 1 \qquad h \in \mathcal{H}, k \in \mathcal{K}_h^t \tag{7.5}$$

therefore it applies that $f_k^h(t) \in \mathcal{K}_h^t$. Therefore, the relations in (7.1)-(7.3) can be reformulated as follows.

$$x_a(t) = \sum_{h \in \mathcal{H}} \sum_{k \in \mathcal{K}_h^t} f_k^h(t) \delta_{ak}^h, \qquad a \in \mathcal{A} \quad (7.6)$$

$$q_h = \sum_{k \in \mathcal{K}_h^t} f_k^h(t) \qquad h \in \mathcal{H} \quad (7.7)$$

$$f_k^h(t) \geq 0 \qquad h \in \mathcal{H}, k \in \mathcal{K}_h^t \quad (7.8)$$

where a link-path incidence matrix element $\delta_{ak}^h = 1$ if link $a$ is part of path $k \in \mathcal{K}_h^t$.

Let then consider the case of separable, differentiable, monotonically increasing link cost functions $\mathbf{c}\big(\mathbf{x}(t)\big) = \{c_a\big(x_a(t)\big), a \in A\}$, where $c_a\big(x_a(t)\big)$ is the travel cost that users experience at day $t$ by travelling through link $a$. Regarding instead the paths, we introduce for each OD pair a matrix of relative-costs $\mathbf{C}^h(t) = \{C_{ks}^h(t), h \in \mathcal{H}, k, s \in K_h^t\}$, where each element $C_{ks}^h(t)$ represents how users travelling on path $k \in \mathcal{K}_h^t$ perceive the costs of path $s \in \mathcal{K}_h^t$ at day step $t$. The relative-cost $C_{ks}^h(t)$ is defined by the sum of two cost components:

$$C_{ks}^h(t) = A_s^h(t) + S_{ks}^h(t) \qquad h \in \mathcal{H}, k \in \mathcal{K}_h^t, s \in \mathcal{K}_h^t \quad (7.9)$$

where

$$A_s^h(t) = \sum_{a \in \mathcal{A}} c_a\big(x_a(t)\big)\delta_{as}^h \qquad h \in \mathcal{H}, a \in \mathcal{A}, s \in K_h^t \quad (7.10)$$

$$S_{ks}^h(t) = \frac{\Psi}{T_s^h(t)} \frac{\sum_{a \in \mathcal{A}} \delta_{ak}^h (1 - \delta_{as}^h) l_a}{\sum_{a \in \mathcal{A}} (\delta_{ak}^h l_a)} \qquad h \in \mathcal{H}, a \in \mathcal{A}, s \in \mathcal{K}_h^t, k \in \mathcal{K}_h^t \quad (7.11)$$

Equation (7.10) allows to compute the actual travel cost of route $s$ as the sum of the actual travel costs of each link within it. On the other hand, (7.11) is an additional cost component representing users' reluctance to change habits. $S_{ks}^h(t)$ denotes the switching cost of users currently on path $k$ willing to change their choice in favor of path $s$, where $l_a$ represents a length measure of link $a$, while $\Psi \geq 0$ determines the weight of this cost component compared with $A_s^h(t)$. As can be seen in (7.11), such cost depends on the overlapping percentage between the two paths, i.e. how much they share some portions of the network. The more the two paths overlap, the lower the switching cost becomes, reaching zero if $k = s$, i.e.:

$$S^h_{kk}(t) = \frac{\Psi}{T^h_s(t)} \frac{\sum_{a\in\mathcal{A}} \delta^h_{ak}(1-\delta^h_{ak})l_a}{\sum_{a\in\mathcal{A}}(\delta^h_{ak}l_a)} = 0 \qquad h \in \mathcal{H}, k \in \mathcal{K}^t_h \quad (7.12)$$

since $\delta^h_{ak}$ is binary, implying that $\delta^h_{ak}(1-\delta^h_{ak}) = 0$.

This leads to the following relation:

$$A^h_k(t) = C^h_{kk}(t) \qquad h \in \mathcal{H}, k \in \mathcal{K}^t_h \quad (7.13)$$

It is worth noting that if $l_a = l, \forall a \in \mathcal{A}$, the switching cost depends purely on the number of shared links between the two paths regardless of their length.

Referring again to (7.11), $T^h_s(t)$ is responsible for the evolution of the switching cost over time. The main idea is that, when a route is significantly used by users of an OD pair, such a route becomes "familiar". Once a route is familiar to users, the perceived costs of switching to it decreases over time until they vanish. On the contrary, a route not significantly used by users is defined as "unfamiliar". In this case the cost of switching from other routes to it does not decrease over time. One way to represent this process is as follows:

$$T^h_s(t) = \begin{cases} t - t^h_s & \text{if } t > t^h_s \\ 1 & \text{otherwise} \end{cases} \quad (7.14)$$

where $t^h_s$ is the day after which the route can be considered familiar. It must hold that $f^h_s(t^h_s) \geq \psi q_h \wedge f^h_s(t) < \psi q_h$, $\forall t < t^h_s$, with $\psi \in [0,1]$. In other words, $T^h_s(t)$ begins to increase when path $s$ starts to be used by a consistent percentage of users, represented by parameter $\psi$. This implies that, for all familiar paths, the switching cost fades to zero over time resulting in $C^h_{ks}(t) \to A^h_s(t)$. That is, users consider only actual path costs in their assessments. It is worth noting, however, that $t^h_s$ does not necessarily exist for each path of pair $h$. In that case, $s$ is still considered unfamiliar and as a consequence $T^h_s(t) = 1$ indefinitely.

Equations (7.9)-(7.13) imply that the relative path cost matrices are square matrices, having the actual costs on the diagonal as follows:

$$\mathbf{C}^h(t) = \begin{bmatrix} A_1^h(t) & \dots & C_{1p}^h(t) \\ \vdots & \ddots & \vdots \\ C_{p1}^h(t) & \dots & A_p^h(t) \end{bmatrix} \tag{7.15}$$

where $p$ is the cardinality of set $\mathcal{K}_h^t$.

### 7.1.1 System dynamics

Making use of the notation and definitions exposed so far, the path flow dynamics can be expressed as:

$$f_k^h(t+1) = f_k^h(t) + L^h(t) \sum_{s \in \mathcal{K}_h} \left[ f_s^h(t) D_{sk}^h(t) - f_k^h(t) D_{ks}^h(t) \right] \qquad h \in \mathcal{H}, k \in \mathcal{K}_h^t \tag{7.16}$$

where:

$$D_{ks}^h(t) = \frac{[A_k^h(t) - C_{ks}^h(t)]_+}{\sum_{\imath \in \mathcal{K}_h} \sum_{\jmath \in \mathcal{K}_h} [A_\imath^h(t) - C_{\imath\jmath}^h(t)]_+ + M} \qquad k \in \mathcal{K}_h^t, s \in \mathcal{K}_h^t \tag{7.17}$$

$$L^h(t) = e^{\phi[\bar{C}^h(t) - E^h(t-1)]_-} \qquad h \in \mathcal{H} \tag{7.18}$$

with projection operators $[\cdot]_+$ and $[\cdot]_-$ defined as $[x]_+ = \max\{x, 0\}$ and $[x]_- = \min\{x, 0\}$.

Referring to (7.17), $D_{ks}^h(t)$ denotes a cost-based swap-rate between path $k$ and path $s$ of OD pair $h$ at day $t$. It represents the percentage of flow $f_k^h(t)$ which will swap towards $s$ on the next day. According to the definition of projection operators, the equation implies that, if the travel cost of route $k$ is higher than the one of route $s$, then a fraction $D_{ks}^h(t)$ of the users will transfer to route $s$. The swapping percentage is equal to the cost difference between $k$ and $s$, normalized over the the total travel cost variation for OD pair $h$ plus $M$, a reluctance coefficient, which represents the insensitivity of users to travel cost differences.

Referring to (7.18), $L^h(t)$ further scales the flows of the pair $h$ that will actually make the shift. This rate encapsulates users' myopia and subsequent optimism when they are facing an improving situation. On a day $t$, the more the average travel cost $\bar{C}^h(t)$ experienced by the users has decreased compared to the historic average travel cost $E^h(t-1)$ experienced up to the previous day, the lower the percentage of those who will decide to change route again on day $t+1$. The $\phi$ coefficient alters the intensity of this behavior. As defined in (7.19), the average cost for pair $h$ is obtained by averaging the actual path costs over the relative flows. As time passes, these costs are embedded within the exponential smoothing defined is (7.20), representing users' past experiences, with $\xi \in (0,1]$.

$$\bar{C}^h(t) = \frac{1}{q^h(t)} \sum_{k \in \mathcal{K}_h^t} f_k^h(t) \cdot A_k^h(t) \qquad h \in \mathcal{H} \quad (7.19)$$

$$E^h(t) = \begin{cases} \bar{C}^h(t_0) & \text{if } t = t_0 \\ \xi \bar{C}^h(t) + (1-\xi) E^h(t-1) & \text{otherwise} \end{cases} \qquad h \in \mathcal{H} \quad (7.20)$$

Figure 7.1 helps to better understand how users' myopic behavior is implemented and the impact of $\phi$ on it. On the x-axis we find the value, at day $t$, of the average cost $\bar{C}^h(t)$, while on the y-axis the corresponding value of $L^h(t)$ is reported. By fixing $\phi$, if $\bar{C}^h(t) \geq E^h(t-1)$, i.e. if the costs experienced by the users of pair $h$ on day $t$ are on average higher than those experienced in the past, then the percentage of users shifting is the highest and then $L^h(t) = 1$. On the contrary, if $\bar{C}^h(t) < E^h(t-1)$ it results that $L^h(t) < 1$. This implies that if users, given their current choice configuration, have experienced a significant drop in travel costs, they are more likely to replicate the same choices on the following day. In other words, a higher proportion of them will decide not to switch routes even if there are better opportunities. The greater the coefficient $\phi$, the less users, given the same value of $\bar{C}^h(t) - E^h(t-1)$, will choose to switch to other paths.

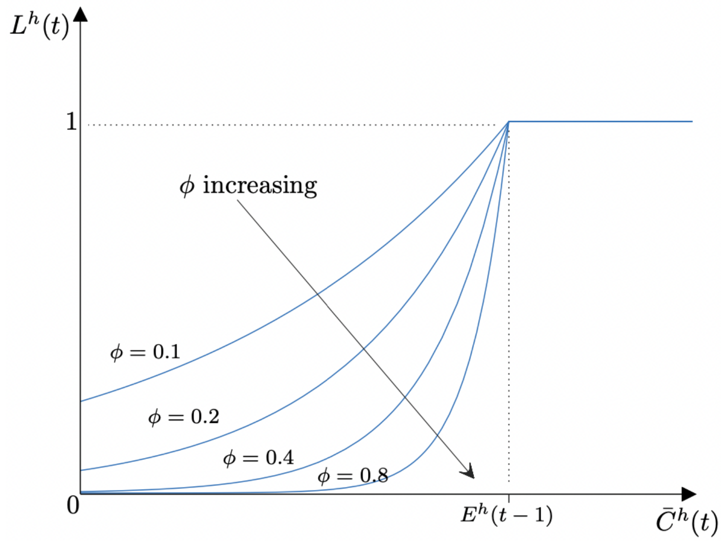



FIGURE 7.1: The influence of $\phi$ over $L^h(t)$

According again to Fig. 7.1 , it can be observed that, when $\phi \to \infty$, the function $L^h(t)$ tends to a step form in the proximity of $E^h(t-1)$. Users are extremely myopic and, as soon as they experience a drop in travel time, they replicate the same choice the next day until circumstances change. On the contrary, $\phi = 0$ represents users who are in no way affected by this behavior. It can be noticed that, if travel costs no longer exhibit significant variations as time passes, it holds that $|E^h(t-1) - \bar{C}^h(t)| \to 0$ by the definition of exponential smoothing. Therefore, $L^h(t) \to 1$ allowing flows to better arrange on the network. Finally, it is worth noting that, by applying (7.18), $L^h(t) > 0$ always holds.

Going back to the dynamics in (7.16) and considering two paths $k$ and $s$, every day the percentage of flow $f_k^h(t)$ shifting from path $k$ towards $s$ or on the contrary the percentage of flow $f_s^h(t)$ shifting from $s$ to $k$ depend on the products $L^h(t)D_{ks}^h(t)$ and $L^h(t)D_{sk}^h(t)$ respectively. However, it is worth noting that if $D_{ks}^h(t) > 0 \implies D_{sk}^h(t) = 0, \forall s, k \in \mathcal{K}_h^t$, and vice versa. Then, keeping in mind (7.13), it is possible to state the following proposition.

**Proposition 1.** *The exchange of flows between two paths of the same OD pair $h \in \mathcal{H}$ is at most unidirectional.*

*Proof.* Let us consider the case in which it is convenient for users on path $k$ to switch to path $s$, with $k, s \in \mathcal{H}$, which is equivalent to $D^h_{ks}(t) > 0$. By considering (7.17) and by applying (7.9)-(7.13) and the definition of projection operators $[\cdot]_+$, it follows that:

$$\begin{aligned} 0 &< C^h_{kk}(t) - C^h_{ks}(t) = C^h_{kk}(t) - A^h_s(t) - S^h_{ks}(t) \\ &< C^h_{kk}(t) + S^h_{sk}(t) - C^h_{ss}(t) \\ &= A^h_k(t) + S^h_{sk}(t) - C^h_{ss}(t) \\ &= C^h_{sk}(t) - C^h_{ss}(t) \end{aligned} \tag{7.21}$$

leading to:

$$[C^h_{ss}(t) - C^h_{sk}(t)]_+ = 0 \implies D^h_{sk}(t) = 0 \tag{7.22}$$

□

This result has an important implication, namely the elements $D^h_{ks}(t)$ do not only represent the swap rate but also the direction of the trajectory at day $t$ of the dynamical system defined in (7.16).

It is also important that the swapping process respects flow conservation. It is therefore possible to state the following proposition.

**Proposition 2.** *The dynamic model in (7.16) ensures the flow conservation.*

*Proof.* Considering each OD pair $h \in \mathcal{H}$, it follows from (7.16) that:

$$\begin{aligned} &\sum_{k \in \mathcal{K}^t_h} f^h_k(t+1) - \sum_{k \in \mathcal{K}^t_h} f^h_k(t) = \\ &= L^h(t) \left[ \sum_{k \in \mathcal{K}^t_h} \sum_{s \in \mathcal{K}^t_h} f^h_s(t) D^h_{sk}(t) - \sum_{k \in \mathcal{K}^t_h} \sum_{s \in \mathcal{K}^t_h} f^h_k(t) D^h_{ks}(t) \right] = \\ &= L^h(t) \left[ \sum_{k \in \mathcal{K}^t_h} \sum_{s \in \mathcal{K}^t_h} f^h_s(t) D^h_{sk}(t) - \sum_{k \in \mathcal{K}^t_h} \sum_{s \in \mathcal{K}^t_h} f^h_s(t) D^h_{sk}(t) \right] = 0 \end{aligned} \tag{7.23}$$

and this proves the flow conservation. □

### 7.1.2 Results on the stationary point

In this section we show that the attitudes described in Section 7.1.1 imply a Bounded Rationality route choice behavior, i.e. the stationary state corresponds to a BRUE. The definition of BRUE via indifference band has been covered in detail in section 4.2 and it is reported briefly here making use of the notation introduced in this chapter.

**Definition 8.** *A path flow* $\mathbf{f}(t) \in \Theta$ *is a BRUE flow pattern if the following holds:*

$$f_k^h(t) > 0 \implies A_k^h(t) \leq \pi^h + \epsilon^h \qquad k \in \mathcal{K}_h^t, h \in \mathcal{H} \quad (7.24)$$

*where* $A_k^h(t)$ *is the actual travel cost on path* $k \in \mathcal{K}_h^t$ *and* $\pi^h$ *is the minimum travel path cost between OD pair* $h \in \mathcal{H}$ *given the flow pattern* $\mathbf{f}^{(t)}$*, while* $\epsilon^h \geq 0$ *is the bounded rationality threshold of users of OD pair* $h \in \mathcal{H}$*.*

The definition of BRUE implies that, at the equilibrium, the travel cost of all routes used by the users of an OD pair is not larger than the absolute minimum plus some margin. The reason for this is that the users are assumed to be indifferent to subtle variations in travel times. Definition 8 has an important implication, namely a BRUE equilibrium is usually not unique. In fact, there may be different patterns of flows that satisfy the condition.

On the contrary, when the threshold falls to zero, the definition reverts to the classical UE, i.e.:

$$f_k^h(t) > 0 \implies A_k^h(t) \leq \pi^h \qquad k \in \mathcal{K}_h^t, h \in \mathcal{H} \quad (7.25)$$

which implies that, for all actually used paths, the travel costs must be equal to the minimum one. It is then worth noting that the UE actually falls within the BRUE set of multiple equilibria.

The following theorem can now be formulated.

**Theorem 1.** *If* $\mathbf{f}(t)$ *is a stationary point, then* $\mathbf{f}(t)$ *is a BRUE.*

*Proof.* Let $\mathbf{f}^t$ be a stationary point for the system. Then, by definition, the following equality must hold:

$$\begin{aligned}
0 &= \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^t} C_{kk}^h \left[f_k^h(t+1) - f_k^h(t)\right] \\
&= \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^t} C_{kk}^h(t) L^h(t) \sum_{s\in\mathcal{K}_h^t} \left[f_s^h(t) D_{sk}^h(t) - f_k^h(t) D_{ks}^h(t)\right] \\
&= \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^t} \sum_{s\in\mathcal{K}_h^t} C_{kk}^h(t) L^h(t) f_s^h(t) D_{sk}^h(t) \\
&\qquad - \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^t} \sum_{s\in\mathcal{K}_h^t} C_{kk}^h(t) L^h(t) f_k^h(t) D_{ks}^h(t) \\
&= \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^t} \sum_{s\in\mathcal{K}_h^t} C_{ss}^h(t) L^h(t) f_k^h(t) D_{ks}^h(t) \\
&\qquad - \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^t} \sum_{s\in\mathcal{K}_h^t} C_{kk}^h(t) L^h(t) f_k^h(t) D_{ks}^h(t) \\
&= -\sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^t} \sum_{s\in\mathcal{K}_h^t} L^h(t) f_k^h(t) D_{ks}^h(t) \left(C_{kk}^h(t) - C_{ss}^h(t)\right)
\end{aligned} \tag{7.26}$$

It can be noticed that each element of the sum is non-negative. In fact by considering (7.8) and (7.18), we can see that $L^h(t) > 0$ and $f_k^h(t) \geq 0$ respectively. Therefore once considered (7.17), $D_{ks}^h(t) \geq 0$ holds by the definitions of projector operator. The only term apparently able to change in sign is $C_{kk}^h(t) - C_{ss}^h(t)$. Considering (7.9)-(7.11) and the associated result in (7.13), it can be observed that, whenever this happens, the projection operator term is equal to zero, as shown below:

$$\begin{aligned}
0 > C_{kk}^h(t) - C_{ss}^h(t) &> C_{kk}^h(t) - C_{ss}^h(t) - S_{ks}^h(t) \\
&= C_{kk}^h(t) - C_{ks}^h(t) \Rightarrow [C_{kk}^h(t) - C_{ks}^h(t)]_+ = 0
\end{aligned} \tag{7.27}$$

which implies that $D_{ks}^h(t) = 0$. Thus each element must be individually equal to zero. This is true exclusively when $f_k^h(t) = 0$ or when $D_{ks}^h(t) = 0$. In fact, under similar considerations as in (7.27), if the projection operator is strictly greater than zero, then $C_{kk}^h(t) - C_{ss}^h(t) > 0$ necessarily.
It is therefore possible to state the following:

$$f_k^h(t) > 0 \implies D_{ks}^h(t) = 0 \qquad k, s \in \mathcal{K}_h^t \tag{7.28}$$

implying that $[C^h_{kk}(t) - C^h_{ks}(t)]_+ = 0$. By applying once again the definition of the projection operator, $D^h_{ks}(t) = 0$ only when $C^h_{kk}(t) \leq C^h_{ss}(t) + S^h_{ks}(t)$, i.e.

$$A^h_k(t) \leq A^h_s(t) + S^h_{ks}(t) \qquad k, s \in \mathcal{K}^t_h \quad (7.29)$$

Equation (7.29) must hold for every used path $k$ with respect to any other path $s$, whether used or not. Chosen $q$ such that $A^h_q(t) = \min_{s \in \mathcal{K}^t_h}\{A^h_s(t)\}$ and $p$ such that $S^h_{pq}(t) = \max_{k \in \mathcal{K}^t_h}\{S^h_{kq}(t)\}$, then the following holds.

$$A^h_k(t) \leq A^h_q(t) + S^h_{pq}(t) \qquad p, q, k \in \mathcal{K}^t_h \quad (7.30)$$

Given that the system is stationary at day $t$ and $S^h_{pq}(t)$ is certainly not increasing over time by the definition in (7.11), let us fix $\pi^h = A_q(t)$ and $\epsilon^h(t) = S^h_{pq}(t)$. It should be noted that, since all the switching costs are not increasing functions over time, the relation would remain valid from $t$ on. It is then sufficient to pick any threshold such that $\epsilon^h > \epsilon^h(t)$. Thus, $\mathbf{f}^{(t)}$ is a BRUE. □

In the following, the type of equilibrium that the system actually reaches and how to estimate $\epsilon^h$ in a reasonable way are further analyzed by means of two remarks. In the former, we consider the case in which $T^h_s(t) = 1, \forall t$, i.e. when the switching cost between two paths depends only on the overlapping percentage and does not change over time. In the latter case, instead, we consider the general case with no assumptions on $T^h_s(t)$.

**Remark 1.** *If the switching costs are fixed* $\left(S^h_{ks} = \Psi \frac{\sum_{a \in \mathcal{A}} \delta^h_{ak}(1-\delta^h_{as}) l_a}{\sum_{a \in \mathcal{A}}(\delta^h_{ak} l_a)}\right)$*, then the stationary point is a BRUE, where* $\epsilon^h = \max_{ks}\{S^h_{ks}\}$.

*Proof.* The result of Theorem 1 holds even if the switching costs are constant. It is sufficient to require that $T^h_s(t) = 1$ indefinitely for every path $s$ of every OD pair $h$. Having indicated with $q$ the path with the minimum objective cost for the pair $h$ and given $S^h_{pq} = \max_k\{S^h_{kq}\}$, the following inequality holds:

$$S^h_{pq} \leq \max_{ks}\{S^h_{ks}\} \tag{7.31}$$

where the term on the right is the maximum switching cost among all paths of pair $h$. Rather intuitively, the largest switching cost with respect to the cheapest path cannot be greater than the largest switching cost regardless of the path for which it is calculated. This allows us to define the threshold as follows.

$$\epsilon^h = \max_{ks}\{S^h_{ks}\} \tag{7.32}$$

The switching costs can be rewritten by extracting the overlapping percentage as:

$$S^h_{ks} = \Psi\Big(1 - \frac{\sum_{a\in\mathcal{A}}(\delta^h_{ak}\delta^h_{as}l_a)}{\sum_{a\in\mathcal{A}}(\delta_{ak}l_a)}\Big) \tag{7.33}$$

For the sake of readability let us define the overlapping percentage as $O^h_{ks} = \frac{\sum_{a\in\mathcal{A}}(\delta^h_{ak}\delta^h_{as}l_a)}{\sum_{a\in\mathcal{A}}(\delta_{ak}l_a)}$, then we can conclude that (7.32) is equivalent to:

$$\epsilon^h = \Psi \cdot (1 - \min_{ks}\{O^h_{ks}\}) \tag{7.34}$$

The result in (7.34) states that we can estimate in advance a maximum threshold for pair $h$ taking into account the swap coefficient $\Psi$ and the two least overlapping paths for pair $h$. □

**Remark 2.** *If the switch costs are time-dependent* $\Big(S^h_{ks}(t) = \frac{\Psi}{T^h_s(t)} \frac{\sum_{a\in\mathcal{A}}\delta^h_{ak}(1-\delta^h_{as})l_a}{\sum_{a\in\mathcal{A}}(\delta^h_{ak}l_a)}\Big)$, *then the result of Remark 1 represents an upper bound for thresholds* $\epsilon^h(t)$.

*Proof.* The $S^h_{ks}(t)$ are non-increasing functions, moreover, it can be noted that $S^h_{ks}(t_0) = S^h_{ks}$ where $S^h_{ks}(t_0)$ is the switching cost for swapping from $k$ to $s$ at $t_0$ and $S^h_{ks}$ is a time-independent switching cost as defined in Remark

1. Then the following applies:

$$S^h_{ks}(t) \leq S^h_{ks} \qquad \forall t > t_0,\ s \in \mathcal{K}^t_h, k \in \mathcal{K}^t_h \quad (7.35)$$

Consequently the following also holds:

$$\max_{ks} S^h_{ks}(t) \leq \max_{ks} S^h_{ks} \qquad \forall t > t_0,\ s \in \mathcal{K}^t_h, k \in \mathcal{K}^t_h \quad (7.36)$$

Finally, by fixing $\epsilon^h(t) = \max_{ks} S^h_{ks}(t)$ and $\epsilon^h = \max_{ks} S^h_{k,s}$, it follows that:

$$\epsilon^h(t) \leq \epsilon^h \quad (7.37)$$

□

Remark 2 states that, since the switching costs cannot increase over time, the threshold $\epsilon^h(t)$ will still be no larger than the one defined in (7.32)-(7.34).

To summarize, as previously defined in (7.14), at a given day $t$ the paths can potentially be divided into two sets: paths labeled as familiar and paths labeled as unfamiliar. For the former, the switching costs decrease over time, while for the latter they remain constant. Given $\mathbf{f}(t)$ as a stationary point of the system, there are two possible scenarios:

1. if the shortest path for pair $h$ ends up being a familiar one, then $\epsilon^h(t) \to 0$. If this happens for every pair, then the system converges to an UE;

2. if, on the contrary, the shortest path is not sufficiently used and therefore unfamiliar, then $\epsilon^h(t)$ does not fade to zero. If this is the case for at least one pair, then the system does not converge to an UE. In fact, said $q$ the shortest path for pair $h$, (7.30) must hold. In this case, however, $S^h_{pq}(t)$ does not fade to zero over time and so $\mathbf{f}(t)$ remains a proper BRUE.

In conclusion, the system, once stationary, is in a BRUE with an upper bounded threshold as demonstrated in Remark 2. If the shortest paths are also familiar for all OD pairs, then all thresholds fade to zero and $\mathbf{f}(t)$ is an UE.

## 7.2 Model Implementation

In this section a viable implementation of the proportional adjustment process described in this chapter is discussed with a particular attention on the initialization phase. Since the work by He, Guo, and Liu (2010), the initialization phase is known to be critical for deterministic path-based DTD models.

In order to determine the initial state of the network, and assuming that the boundary conditions have remained unaltered for a sufficiently long time to allow the dynamics of the system to stabilize, it is convenient to make use of traditional static assignment models. It is well known that the solution of deterministic models, under mild conditions, is unique with respect to link flows but not with respect to path flows. In other words, the pattern of link flows corresponding to the computed equilibrium is unique but at the same time it can be associated with multiple path flow patterns. The choice of the initial path flow pattern is therefore discretional. This would not be a problem if different initial states implied identical dynamics, but this is not necessarily true. We can therefore employ a path flow estimation technique of the kind described by Larsson et al. (2001). In particular, in this work, we make use of the entropy maximization approach proposed by Rossi, McNeil, and Hendrickson (1989), which we briefly outline here.

Let us consider every single trip as a non-fungible entity, i.e. with its own identity and therefore uniquely identifiable. Given a path flow pattern $\mathbf{f}$, we define as "state" a specific allocation of each trip consistent with $\mathbf{f}$. Assuming that each state is equally like, then the most likely flow pattern $\mathbf{f}$ will be the one that allows the highest number of admissible states, i.e. the one corresponding to the highest level of disorder. The resulting entropy function to be maximized as follows:

$$\max g(\mathbf{f}) = \prod_h \frac{q_h!}{\prod_k f_k^h!} \tag{7.38}$$

After applying few algebraic manipulations and the Stirling's formula (see Rossi, McNeil, and Hendrickson (1989) for details), $\max g(\cdot)$ is approximated by $\min h(\cdot)$ and the resulting optimization problem becomes:

**Problem 8.**

$$\textit{min}\ \ h(\boldsymbol{f}) = \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^{t_0}} f_k^h \cdot \ln(f_k^h) \tag{7.39}$$

*subject to*

$$\sum_{k\in\mathcal{K}_h^{t_0}} f_k^h = q_h \qquad h \in \mathcal{H} \tag{7.40}$$

$$x_a^{\mathbf{UE}} = \sum_{h\in\mathcal{H}} \sum_{k\in\mathcal{K}_h^{t_0}} f_k^h \cdot \delta_{ak}^h \qquad a \in \mathcal{A} \tag{7.41}$$

where $\mathcal{K}_h^{t_0}$ is the path set for OD pair $h$ when $t = t_0$, i.e. at the equilibrium before the occurrence of the disruption.

Note that function $h(\mathbf{f})$ is strictly convex. Moreover, let $\mathbf{f}^{\mathbf{UE}}$ be an optimal flow pattern for the optimization problem. Constraints (7.40)-(7.41) imply that $\mathbf{f}^{\mathbf{UE}}$ belongs to $\Theta$ given $\mathbf{x}^{\mathbf{UE}}$, the link flow pattern at the equilibrium. The link-path feasible set $\Theta$ is convex, therefore $\mathbf{f}^{\mathbf{UE}}$ is unique. Finally, to initialize the DTD dynamics, it is sufficient to set:

$$\mathbf{f}(t_0) = \mathbf{f}^{\mathbf{UE}} \tag{7.42}$$

Once the initial flow pattern is computed, the network is perturbed. In the numerical examples discussed in Section 7.3, the perturbation is represented by a network disruption resulting in users being prevented from using one or more links. Following the removal of a link, all directly impacted flows are reassigned on the cheapest routes according to (7.9)-(7.11). Thus, $\mathbf{f}(t_1)$ is obtained. At this point the dynamics can take place. At each iteration, the following quantities are updated: path costs matrices $\mathbf{C}^h(t)$ are computed according to (7.9)-(7.11); cost-based swap rates are the updated by (7.17); average pair travel costs $\bar{C}^h(t)$ and the exponential smoothing $E^h(t)$ are estimated by applying (7.19)-(7.20); finally myopia-based swap rates are computed based on (7.18). It is now possible to update the system's dynamics in (7.16) thus obtaining $\mathbf{f}(t+1)$. Then, if a certain convergence criteria is met, the algorithm stops, otherwise the time variable $t$ is updated and the process continues.

A viable implementation of the algorithm is summarized as follows:

**Input:** Network, Demand, Disrupted links
$\mathbf{x}^{\mathbf{UE}} \longleftarrow$ UE traffic assignment;
$\mathbf{f}(t_0) \longleftarrow$ Entropy Maximization by (7.39)-(7.41);
$\mathbf{f}(t_1) \longleftarrow$ Disruption;
**repeat**
  $\mathbf{C}^h(t) \longleftarrow$ relative path cost matrices update by (7.9)-(7.11);
  $\mathbf{D}^h(t) \longleftarrow$ swap-rates update by (7.17);
  $\bar{C}^h(t)$, $E^h(t) \longleftarrow$ average costs and exponential smoothing update by (7.19)-(7.20);
  $L^h(t) \longleftarrow$ dynamics rate update by (7.18);
  $\mathbf{f}(t+1) \leftarrow$ next path flow vector estimation by (7.16)
**until** *convergenze is met*;

From a computational point of view, updating the dynamic process described by (7.16) requires the execution, at each iteration, of simple algebraic computations and for this reason, in line with most DTD models in the literature, the performance scales well even on large networks. It should also be noted that the main goal of this type of models is to mimic the behavior of an associated real network and, for this reason, the speed of convergence, i.e. number of iterations, should be evaluated only with respect to the accuracy of the representation. On the other hand, the initialization phase can be computationally demanding because two optimization problems need to be solved. By solving the well-known Beckmann's Problem (Beckmann, McGuire, and Winsten, 1956), the deterministic User-Equilibrium link flow pattern can be estimated, while solving the Entropy

Maximization problem in (7.39)-(7.41) allows the most likely associated path flow pattern to be univocally determined.

The Frank-Wolfe algorithm (Frank and Wolfe, 1956) has been one of the most widely adopted approaches to solve the deterministic traffic assignment problem. Its strengths are that it requires low memory, no path enumeration is needed, and the implementation procedure is straithforward, characterized by a sequence of all-or-nothing assignments. At the same time, it has a significant drawback. Once in the proximity of the optimum, the algorithm asymptotically converges sub-linearly, since the descent directions tend to become normal to the gradient resulting in a zig-zagging behavior. To overcome these limitations, several link-based algorithms have been proposed to enhance local convergence (Fukushima, 1984; Hearn, Lawphongpanich, and Ventura, 1985; Florian, Guálat, and Spiess, 1987), as well as more recent path-based algorithms (Bar-Gera, 2002; Dial, 2006; Florian, Constantin, and Florian, 2009; Kumar and Peeta, 2010) which have become a suitable alternative given the memory capabilities of modern computers. In particular, bush-based algorithms (Bar-Gera, 2002; Dial, 2006) sequentially decomposing flows by origins exploit the consequence that flows by origins constitute acyclic sub-networks. Flows are transferred from the longest path onto the shortest path, and without cycles the related computations become highly efficient, enabling their use on large-scale networks.

At the same time, the entropy maximization problem described in (7.39)-(7.41), making use of the Stearling's approximation, enables its use on networks of conspicuous size. However, the method is not insensitive to the scale of the problem. Alternatively, Bar-Gera (2006) proposed a primal method that exploits the proportionality condition (Bar-Gera and Boyce, 1999) associated with an entropy maximization problem. Such condition states that, when facing two alternative sections, travelers will distribute equally regardless of their origin and destination. Bar-Gera's algorithm shows good performance even on large-scale networks particularly when link flow are computed by a bush-based algorithm. More recently Kumar and Peeta (2015b), extending the concept of proportionality, compute the single path flow pattern out of the entropy weighted average of all possible path flow patterns. The algorithm is easy to implement and requires little computational effort to generate the solution which makes it an excellent alternative for large-scale networks.

The above mentioned methods, providing the same initial solution, can be employed indifferently within the initialization phase required by the proposed model depending solely on the requirements dictated by the scale of the specific case study.

## 7.3 Numerical Examples

In this section, the proposed model is applied to two example networks. The former (Network 1) is a simple network and the main goal of the tests is to illustrate the different results with different values of the switching cost coefficient $\Psi$. The latter (Network 2) is a larger network and the tests are aimed to show the system behavior in a more complex scenario.

Some metrics are introduced, referred to a generic day $t$, to better describe the performance of the system state:

$$P_a(t) = \frac{c_a(t_0)}{c_a(t)} \qquad a \in \mathcal{A} \quad (7.43)$$

$$P_k^h(t) = \frac{A_k^h(t_0)}{A_k^h(t)} \qquad h \in \mathcal{H}, k \in \mathcal{K}_h^t \quad (7.44)$$

$$P^h(t) = \frac{\bar{C}^h(t_0)}{\bar{C}^h(t)} \qquad h \in \mathcal{H} \quad (7.45)$$

$$P(t) = \frac{\bar{C}(t_0)}{\bar{C}(t)} \quad (7.46)$$

Equations (7.43)-(7.46) define the link performance, the path performance, the average performance for an OD pair, and finally the overall average performance of the system at any given day $t$, respectively. In (7.45), the average costs $\bar{C}^h(t)$ are those defined in (7.19), while the global average costs $\bar{C}(t)$ are, quite similarly, defined as:

$$\bar{C}(t) = \frac{\sum_{h \in \mathcal{H}} \bar{C}^h(t) \cdot q^h}{\sum_{h \in \mathcal{H}} q^h} \qquad h \in \mathcal{H} \quad (7.47)$$

Remembering that $t_0$ represents the pre-disruption setting, it is worth noting that each defined metric identifies the current performance of each component of the system by comparing the current travel costs with those computed in $t_0$. In other words, on a given day $t$, the performance degrades proportionally to travel cost increase at day $t$ (whether computed for a link, a path, an OD pair, or for an entire network) compared with the same cost at day $t_0$. This is a rather standard way of defining the performance of a transportation network. In Zhou, Wang, and Yang (2019) an extensive literature on the topic can be found, with a focus on the metrics that are most used in network resilience evaluations.

### 7.3.1 Network 1

The considered network is shown in Fig. 7.2, it is composed of 7 nodes and 8 links. Each link is characterized by the same performance function, relating the travel cost to the amount of congestion, which is shown below:

$$c_a(x_a) = 10^{-3} \cdot x_a + 10^{-1} \qquad a \in \mathcal{A} \quad (7.48)$$

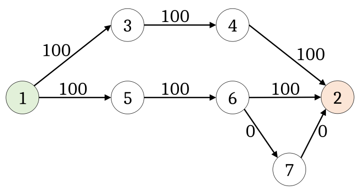
This network presents only one origin-destination pair 1-2. According to the network topology, 3 possible paths join origin node 1 with destination node 2. Table 7.2 shows the incidence relationships between routes and network arcs. The transportation demand associated with the OD pair 1-2 is 200.

| OD pair | Path | Nodes sequence |
|---|---|---|
| (1,2) | 1 | [1,3,4,2] |
| | 2 | [1,5,6,2] |
| | 3 | [1,5,6,7,2] |

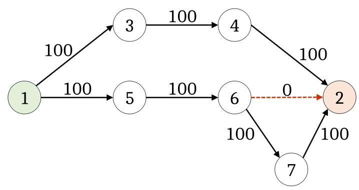
TABLE 7.2: Incidence relations in Network 1

Referring to this network, two different disruptions are tested:

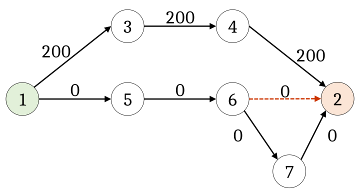
- removal of link 6-2;
- removal of link 1-3.



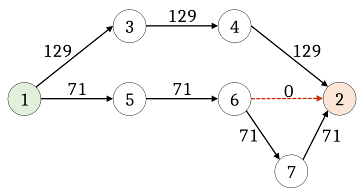

(a) Pre-disruption equilibrium




(b) First flow reassignment with $\Psi = 0.1$



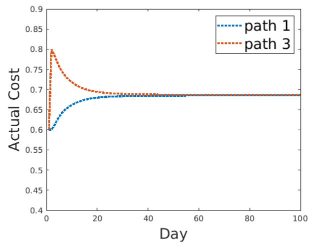

(c) First flow reassignment with $\Psi = 0$



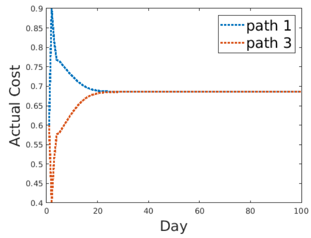

(d) Post-disruption Equilibrium

FIGURE 7.2: Network flows in Network 1 with removal of link 6-2

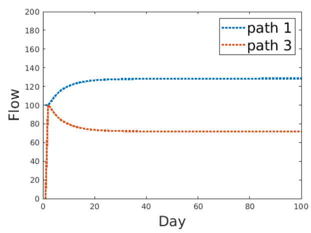
The tests on Network 1 aim to show the evolution of the system after a disruption, depending on the switch coefficient $\Psi$, corresponding to the reluctance of users to use paths topologically dissimilar from those currently used. All the other parameters are kept constant and assume the following values: $\psi = 0.01$, $\phi = 50$, $\xi = 0.6$ and $M = 3$.



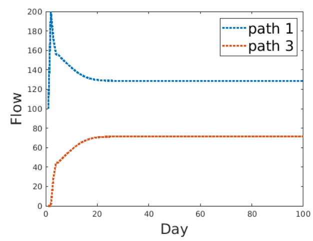

(a) $\Psi = 0.1$



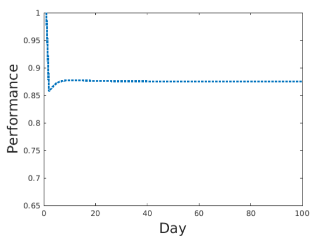

(b) $\Psi = 0$



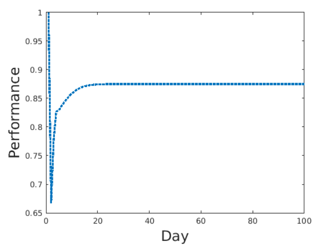

(c) $\Psi = 0.1$




(d) $\Psi = 0$

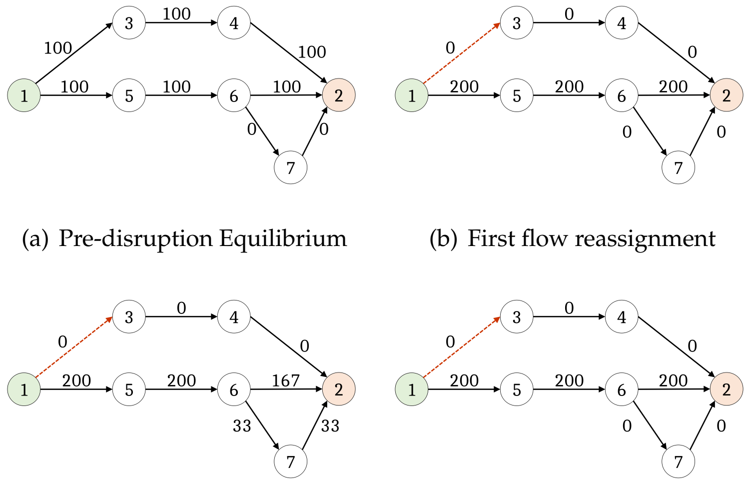



(e) $\Psi = 0.1$




(f) $\Psi = 0$

FIGURE 7.3: Path actual cost, flow, performance evolution in Network 1 with removal of link 6-2




(a) Pre-disruption Equilibrium (b) First flow reassignment

(c) Post-disruption Equilibrium with $\Psi = 0.1$ (d) Post-disruption Equilibrium with $\Psi = 2.7$

FIGURE 7.4: Network flows in Network 1 with removal of link 1-3

Figure 7.2 shows 4 snapshots of the network state taken in 3 different moments of the system evolution after the removal of link 6-2 (the number reported close to each link represents the the flow). In particular, in Fig. 7.2(a) the pre-disruption equilibrium is shown, which does not depend on $\Psi$. Reminding that all the arcs are qualitatively identical, the demand is divided between paths 1 and 2, consisting of 3 links, while path 3, composed of 4 links, is not used at all. After the removal of link 6-2, we see a forced reassignment of flows previously present on path 2: in Fig. 7.2(b), the result of such assignment is shown in case the switch coefficient is set to $\Psi = 0.1$, while Fig. 7.2(c) reports the assignment results with $\Psi = 0$, i.e. when the flows of the users are rearranged on the network based exclusively on actual path costs. As it can be noticed, when $\Psi = 0.1$ all the flow is assigned to path 3, the one sharing a relatively wide portion of network with path 2, now unusable. By adopting this solution, the assignment immediately after the disruption consists of an overall moderate redistribution of flows, reflecting a more conservative user behavior. When $\Psi = 0$, instead, the disruption is followed by a substantial redistribution of flows, now all loaded on path 1. In this case, the users evaluate only the actual path costs, regardless of the degree of similarity between the new chosen path and the abandoned one. Both in case $\Psi = 0.1$ and $\Psi = 0$, the new equilibrium reached by the system is the same, as shown in Fig. 7.2(d). Whether path 3 or path 1 is preferred at first, the system evolves toward the same equilibrium, which is in this specific case a UE. This happens because in both cases the two paths end up being used significantly by users (i.e. they are considered familiar paths) and as described in (7.14) their switch costs $S^1_{31}(t)$ and $S^1_{13}(t)$ fade to 0 over time.

Figure 7.3 shows the dynamics of the system in terms of actual path cost, flows and overall average system performance. In particular, Figs. 7.3(a), 7.3(c) and 7.3(e) show these trends in case $\Psi = 0.1$, while Figs. 7.3(b), 7.3(d) and 7.3(f) report the same variables when $\Psi = 0$. Confirming that the stationary state reached by the system is in fact a UE, in Fig. 7.3(a) and in 7.3(b) it is possible to observe that the actual path costs converge to the same value of about 0.68. However, the way in which the system converges is significantly different in the two cases, with the former showing considerably smaller cost fluctuations.

Again in Fig. 7.3, it is possible to see a similar trend characterizing the path flow evolution. In case $\Psi = 0.1$, from a situation in which the transportation demand is equally distributed on the two remaining paths, path 3 begins to lose flow in favor of path 1 until the equilibrium is reached (see Fig. 7.3(c)). When $\Psi = 0$, instead, path 1 is the one that initially carries all the demand before it redistributes (see Fig. 7.3(d)). These two different dynamics have quite different impacts on the average system performance, as shown in Fig. 7.3(e) and 7.3(f). Even if in both cases the performance level settles around 87% of its original level, the values reached in the worst moment after the disruption differ considerably. In the former case, the system performance does not fall below 85% while in the second case the network performance collapses to about 66% of the original value at the most critical instant. The reason for this is that modeling users who prefers swapping between topologically similar paths results in a much smoother flow redistribution. This, especially right after the disruption, implies less performance loss.

The results obtained considering link 1-3 removal on the same network are shown in Fig. 7.4, reporting again 4 snapshots of the network. Specifically, the traffic pattern after the disruption is depicted in Fig. 7.4(b), while Figs. 7.4(c) and 7.4(d) show the new equilibrium reached by the system in case the switch coefficient is set to 0.1 and 2.7 respectively. As can be seen, they differ significantly. In particular, in the latter case, the system does not evolve anymore after the disruption and the only path to be used is the second one. This happens because the switching cost $S^1_{23}(t)$ is relatively high compared to the actual path cost $A^1_2(t)$ and this means that, even if the cost on path 2 is far from being small, the users are reluctant to use other solutions. More specifically, as long as $A^1_2(t) < S^1_{23}(t)$, cost-based swap rate $D^1_{23}(t) = 0$ by the definition in (7.17).

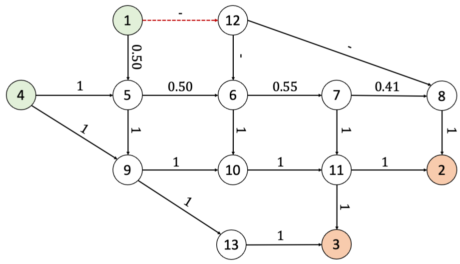
### 7.3.2 Network 2

Network 2 is shown in Fig. 7.5, composed of 19 links, qualitatively identical and with the same performance function reported in (7.48), and 13 nodes. In this network, the origin nodes are 1 and 4 while the destination nodes are 2 and 3. The mobility demand is the same for each of the four OD pairs [1-2, 1-3, 4-2, 4-3] and is equal to 200 units. In this network only one disruption is tested: removal of link 1-12.

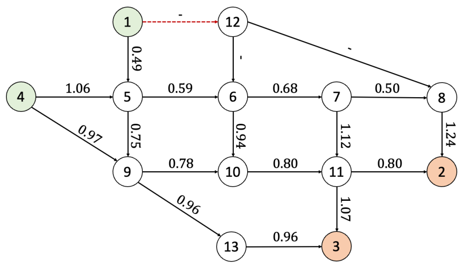
The incidence relationships between paths and links are shown in Table 7.3, where the underlined paths are those unavailable after the disruption. Moreover, the flows in the pre-disruption scenarios are reported.

In Fig. 7.5, the evolution of the network in the initial fifteen days after the disruption is shown. This is the time interval during which the most intense flow rearrangements take place. The performance values, as defined in (7.43), are reported and depicted close to each link.

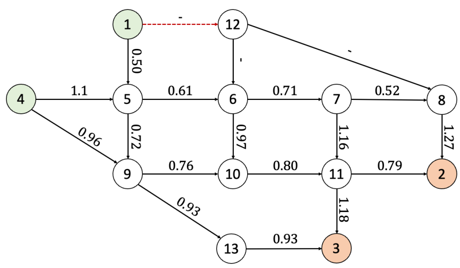
As can be seen in Fig. 7.5(a), once the disruption has taken place, the most stressed links are those in its proximity. As the days go by, as shown in Fig. 7.5(b)-7.5(d), we can see a generalized performance degradation that spreads like a wave throughout the network, while the previously overloaded links recover part of their lost efficiency. As we move away from the disruption site, the magnitude of the disruption decreases.



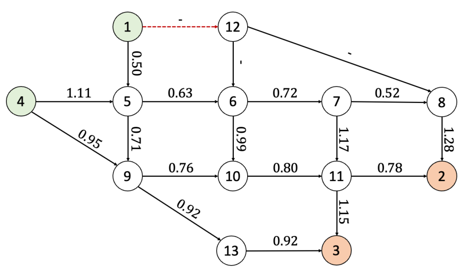

(a) Link Performance day 1




(b) Link Performance day 5




(c) Link Performance day 10




(d) Link Performance day 15

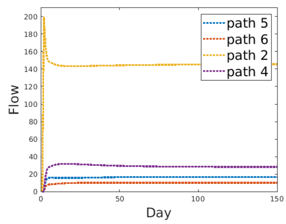
FIGURE 7.5: Link performance in Network 2

| OD pair | Path | Nodes sequence | $\mathbf{f}(t_0)$ | OD pair | Path | Nodes sequence | $f(t_0)$ |
|---|---|---|---|---|---|---|---|
| (1-2) | 1 | [1,12 8,2] | 200 | (4-2) | 1 | [4,9,10,11,2] | 104 |
| | 2 | [1,5,6,7,8,2] | 0 | | 2 | [4,5,6,7,8,2] | 36 |
| | 3 | [1,12,6,7,11,2] | 0 | | 3 | [4,5,9,10,11,2] | 0 |
| | 4 | [1,5,9,10,11,2] | 0 | | 4 | [4,5,6,10,11,2] | 23 |
| | 5 | [1,5,6,10,11,2] | 0 | | 5 | [4,5,6,7,11,2] | 37 |
| | 6 | [1,5,6,7,11,2] | 0 | | | | |
| | 7 | [1,12,6,10,11,2] | 0 | | | | |
| | 8 | [1,12,6,7,8,2] | 0 | | | | |
| (1-3) | 1 | [1,5,9,13,3] | 90 | (4-3) | 1 | [4,9,13,3] | 173 |
| | 2 | [1,12,6,10,11,3] | 18 | | 2 | [4,5,9,13,3] | 7 |
| | 3 | [1,5,6,7,11,3] | 37 | | 3 | [4,9,10,11,3] | 20 |
| | 4 | [1,5,9,10,11,3] | 0 | | 4 | [4,5,6,7,11,3] | 0 |
| | 5 | [1,12,6,7,11,3] | 33 | | 5 | [4,5,9,10,11,3] | 0 |
| | 6 | [1,5,6,10,11,3] | 22 | | 6 | [4,5,6,10,11,3] | 0 |

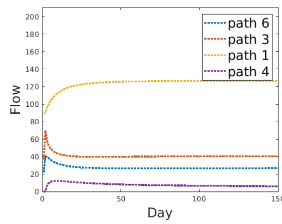
TABLE 7.3: Link-path incidence relationship and pre-disruption assignment in Network 2

It is interesting to note that some links perform better in comparison to the pre-disruption scenario. This is the case of links 8-2, 7-11 and 11-3. Not surprisingly, all those links belong to paths that are no longer available and at the same time the new flow pattern makes less use of them, resulting in a reduction in the level of congestion on these links and, consequently, an increase in performance.

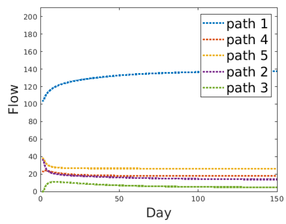
The same network evolution is reflected by looking at the path flow, actual cost and performance trajectories presented, respectively, in Fig. 7.6, 7.7, 7.8. As it is shown in Fig. 7.6(a), path 2 is heavily loaded just after the occurrence of the disruption. Over the following days part of these flows shift to the remaining available paths, as mentioned above, affecting the flow patterns in the remaining links of the network. The OD pair 4-3 shows the smallest flow adjustment before reaching a new stationary state which is consistent with the fact that it is the pair furthest from disruption and whose paths are least affected by flow rearrangement. The new equilibrium, as illustrated in Fig. 7.7, is a proper UE for every OD pair $h$ but with significant differences. For the 1-3 pair, the distribution of demand among the routes is more homogeneous than for the other OD pairs where instead one route in particular is significantly preferred over the others.



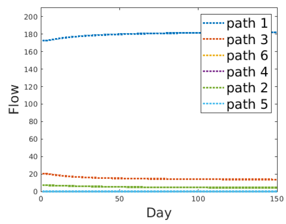

(a) OD 1-2




(b) OD 1-3

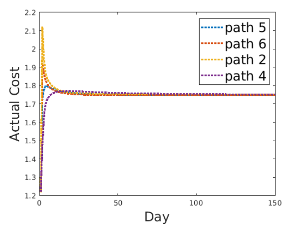



(c) OD 4-2

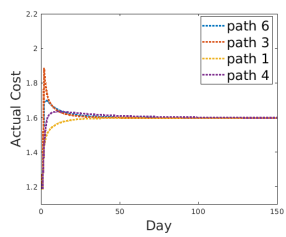



(d) OD 4-3

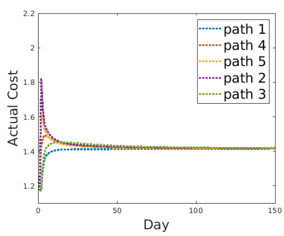
FIGURE 7.6: Path flow evolution in Network 2



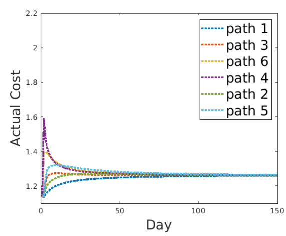

(a) OD 1-2




(b) OD 1-3

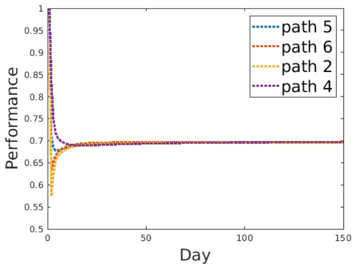



(c) OD 4-2

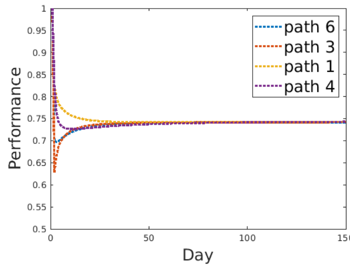



(d) OD 4-3

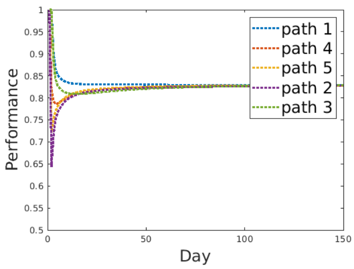
FIGURE 7.7: Path actual cost evolution in Network 2



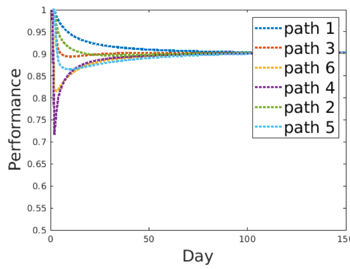

(a) OD 1-2




(b) OD 1-3




(c) OD 4-2




(d) OD 4-3

FIGURE 7.8: Path performance evolution in Network 2

The trend of path performance illustrated in Fig. 7.8 further confirms that the closer a path is to the disruption epicenter, the more it is affected by it. At the critical moment, the performance degradation is in fact maximum for the paths of the OD pair 1-2 (Fig. 7.8(a)), where path 2 is the one suffering most from the new network configuration, with performance reaching 57% of the original value. Regarding OD pairs 1-3 and 4-2 (respectively shown in Fig. 7.8(b) and 7.8(c)), paths 2 and 3 suffer the greatest degradation in performance with 63% and 65% of pre-disruption values respectively. Lastly, the paths of OD pair 4-3 do not experience any significant performance loss (Fig. 7.8(d)). Path 4, the only one showing significant performance degradation, is in fact basically unused, as shown in Fig. 7.6(d).

Finally, the average performance of each OD pair as defined in (7.45) are reported in Fig. 7.9 which once again confirms the tendency whereby the OD pairs close to the site of the disruption are the most affected by it.In particular, the OD pairs 1-2, 1-3 and 4-2 all exhibit, although to different magnitudes, a drastic drop in performance immediately after the disruption, before gradually recovering. The OD 1-2 pair is the one that suffers the most severe performance drop of roughly over 50%. In contrast, the OD 4-3 pair is not only the least affected by the disruption but is also the only one that does not suffer an immediate performance loss. Instead, its performance degrades as the performance of the others improves due to a better redistribution of flows.

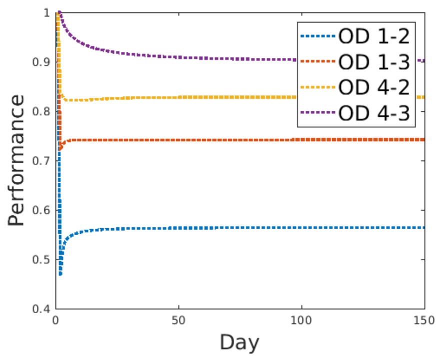



FIGURE 7.9: O-D performance evolution in Network 2

## 7.4 Conclusions

In this chapter a day-to-day discrete-time assignment model based on a proportional switch adjustment process is proposed. The model defines the amount of users who, unsatisfied with their current status, decide for the next day to satisfy their mobility needs by switching paths. The users' route choice behavior makes the routes that have the potential to reduce travel costs preferable. The greater the decrease in costs, the higher the percentage of users who switch. The model incorporates some cognitive biases within the users' route choice process. The first concerns a spatial inertia, whereby users overestimate the cost of potential paths in relation to the level of topological similarity they present with respect to the path they are currently using. The more the path they are currently on differs topologically from a potential candidate, the higher the perceived cost of abandoning the former in favor of the latter. This behavior is incorporated into the cost structure of the paths themselves. In addition to the additive component, related to the actual state of congestion on the network links, there is a second component defined as a switching cost, which are higher if two paths overlap for small portions. The cost of a route is therefore not unique but depends on where the users currently evaluating it are located. The adjustment process describes users who prefer switching between topologically similar paths. In parallel, users also exhibit a myopic behavior. After a switch, if the travel cost experienced is significantly lower than the one they are used to, they will tend to stick with it for the following days even though the conditions may have changed in the meantime. Finally, it is assumed that these biases may fade over time with respect to those routes that the users are more familiar with.

We have shown that the new equilibrium state must necessarily fall into a BRUE, a set of states that, while not necessarily an UE, do not diverge from it more than a certain value which increases when the relative importance given to the topological similarity within the choice process grows. The examples reported show how, depending on the circumstances, the switching cost can significantly affect both the trajectory of the system and the equilibrium it reaches. Moreover, the example applied on the larger network shows how the model can to represent the cascading effects in performance degradation that would be expected in the presence of a disruption. The paths to be more influenced by the disruption tend to be those nearer to it, while those farther away are significantly less affected by the perturbations that involve the network. This model allows, as in this case, to consider the tendency of users to prefer the routes they use most against other solutions potentially superior but somewhat unusual. That said, the proposed model can be also used to represent, appropriately redefining the switching cost component, a variety of other scenarios where it may be necessary to characterize the mobility choice consequences based on different group of users.

# Chapter 8

# Conclusions

The primary objective of the present thesis has been the development of methodologies suitable for representing the macroscopic dynamics that arise when a transportation network is subjected to a disruptive event. Transportation networks are inherently complex systems where techno-social elements interoperate resulting in complex processes. For this reason, assessing the level of vulnerability or resilience of a transportation network based solely on its topology is a partial assessment. Therefore, appropriate models able to dynamically account for the interaction between transportation supply and demand have been developed with particular attention to the assumptions underlying travelers' behavior and the interaction between multiple transportation modalities. To this end, several traffic assignment models have been developed within this thesis.

First, a review of the literature on disruption analysis of transport networks was carried out, and it emerged that assignment models could be a valuable tool allowing a sufficiently accurate representation of the equilibration processes that occur on a network when affected by a disruption without, however, the computational burden resulting from excessively fine modeling of the network. A careful review of traffic assignment models was then conducted, paying particular attention to the various classifications into static, fully dynamic and semi-dynamic models. Based on the review process, it was possible to deduce that semi-dynamic (inter-periodic or day-to-day) traffic assignment models are best suited to represent the equilibration processes that occur when the status quo of the transportation system is altered. For each period, they allow to estimate the percentage of users who given the new network conditions will decide to modify their travel choices. The resulting flow adjustments in turn alter the congestion and travel cost patterns on the links therefore further altering the network conditions under which travelers will once again have to make their choices. This circular feedback process, in the absence of further shocks, tends to stabilize. Different assumptions about users' choice behavior result in system trajectories and equilibria exhibiting different characteristics.

In this regard, an analysis of the literature on bounded rationality has been carried out. This conceptual framework, widely used in economic disciplines, proposes the idea that people may not exhibit perfect rationality within their decision process and that therefore their judgment may be affected by systematic bias due to habit or the cognitive burden involved in the decision process, to name a few.

In light of these considerations, a day-to-day link-based traffic assignment model has initially been proposed which accounts for user habits within the evolution process of a transportation system. The model is characterized by a series of static traffic assignments computed out of a set of varying conditions. The ideas behind the model are twofold. The first is that it is assumed that users are insensitive to excessively minute variations in travel costs. This is represented by an index of users' tolerance towards increasing travel costs. The second idea is that users tend to stick to the same routes they have used in the past. According to the model, they decide to use new solutions only when the increase in travel costs can no longer be tolerated, i.e. when the increase exceeds a certain threshold. When this occurs a path is added to the set of available paths, on which assignments are computed. Simultaneously, it is unreasonable to assume that all flows will instantly dispose optimally on the new set of paths. In other words, despite the increase in travel costs it is assumed that not all users may decide to change routes. As a result, an inertia coefficient is considered to prevent the system from jumping from one configuration to another too rapidly.

When a transportation system experiences a particularly significant disruptive event, the demand response may affect multiple transportation modes. In a complex network, the interdependence between different transport modes can be relevant and critical events occurring on one mode can lead to ripple effects involving other transport modes as well, therefore the estimation of the vulnerability of a transport system as a whole cannot avoid considering this possibility. In light of these considerations, a scenario-based analysis was then proposed where multiple transportation modes are integrated. The transportation system is represented by means of a hyper network, where some elements mirror the actual infrastructure while others are associated with user choices. In the proposed model, these choices concern modal shift. The peculiarity of this approach is that it allows users to plan their trip by possibly recurring to multi-modality in the ways they consider most beneficial. In other words, modal choice is not constrained to a pre-trip phase, but instead the traveler can choose a solution that integrates several modes within the same trip. However, such freedom has been limited within reasonable modal change sequences to avoid modeling solutions may feasible in theory but unlikely in reality. Another peculiarity of the model is to consider a multi-class traffic flow. In particular, passenger flows and freight flows are considered are associated with distinct choice behaviors. The scenario-based analysis therefore consists of two steps. In the first step, two distinct assignment models estimate travel choices for passenger flows and freight flows where the former tend to act selfishly by minimizing their own individual travel costs while the latter instead aspire to minimize a generalized fleet cost. Traveler choices are then consistently routed over the network by means of a discrete-time dynamic flow model that allows for more realistic estimates related to flow propagation and opens up the possibility of defining performance indices that take into account more detailed information such as average speeds or pollutant emissions. The model is then tested under two scenarios, before and after the occurrence of a disruption, and changes in the state of the multi-modal network are analyzed.

The insight related to topological inertia associated with user choice behavior introduced in the first work is extended and further explored in a day-by-day path-based traffic assignment model characterized by a proportional switch adjustment process. Defining a day-to-day dynamics on path flows instead of link flows makes it easier to introduce a wide variety of behavior patterns because users' route choice process can be explicitly represented. A second advantage, is that with the exception of the pre-disruption equilibrium state estimation, the dynamics of the system does not require iterated optimization problem solving. In this work, an idea is advanced that users in their choice process are influenced not only by the current state of the network but also by their mobility habits. Where driven by new conditions, users favor routes that are less expensive but also similar to the one they decided to abandon. Two routes are more similar the more they overlap. This type of approach implies that each path is associated with a cost vector, whose individual elements represent the cost as it is perceived by users currently on a specific other path. Thus, it is shown that this assumption implies representing users who are rationally bounded in their decision-making process. An approach for estimating the associated indifference band is then suggested. The model is then tested on a network affected by a disruption and the effects that users' behavioral assumptions may have on the new equilibrium state are discussed.

The research presented in this manuscript can be developed in three directions. First, the demand elasticity can be considered within the analysis. It is reasonable to assume that following major disruptions, transportation demand may drop. A viable approach, similarly to the one followed in chapter 6, would be to extend the actual network to an hyper-network that would take into account any unfulfilled demand. The extension of the presented day-to-day assignment models on an hyper-network is strait forward.

A second direction might be the integration of the model presented in chapter 7 within the assignment process of the multi-modal framework discussed in chapter 6. Once the assignment model is adapted to consider shorter periods within a single day, it could also be used for more detailed estimation of the immediate effects on traffic flows following a disruption.

Finally, the analysis approach presented can be employed in a case study for estimating the vulnerability of a real transportation network. It would be possible to significantly reduce the computational effort of the analysis by pre-selecting via topological metrics a subgroup of network components considered most critical. Then on these elements a stress test can be performed using the approaches presented in this paper in order to model the reaction of the system.

# Bibliography

Abdelfatah A. S. and Mahmassani H. S. (2001). "A simulation-based signal optimization algorithm within a dynamic traffic assignment framework". In: *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 01TH8585)*. IEEE, pp. 428–433.

Adams T. M., Bekkem K. R., and Toledo-Durán E. J. (2012). "Freight resilience measures". In: *Journal of Transportation Engineering* 138.11, pp. 1403–1409.

Adjetey-Bahun K., Birregah B., Châtelet E., and Planchet J.-L. (2016). "A model to quantify the resilience of mass railway transportation systems". In: *Reliability Engineering & System Safety* 153, pp. 1–14.

Agrawal B. B. and Ziliaskopoulos A. (2006). "Shipper–carrier dynamic freight assignment model using a variational inequality approach". In: *Transportation research record* 1966.1, pp. 60–70.

Agrawal S., Zheng H., Peeta S., and Kumar A. (2016). "Routing aspects of electric vehicle drivers and their effects on network performance". In: *Transportation Research Part D: Transport and Environment* 46, pp. 246–266.

Ahmed K. I. (1999). "Modeling drivers' acceleration and lane changing behavior". PhD thesis. Massachusetts Institute of Technology.

Alam M. J., Habib M. A., and Quigley K. (2017). "Vulnerability in transport network during critical infrastructure renewal: Lessons learned from a dynamic traffic microsimulation model". In: *Procedia Computer Science* 109, pp. 616–623.

Almond J. (1965). "Traffic assignment to a road network". In: *Traffic Engineering and Control* 6.10, pp. 616–617.

Almond J. (1967). "Traffic assignment with flow-dependent journey times". In: *Vehicular traffic science*, pp. 222–234.

Almotahari A. and Yazici A. (2020). "Impact of topology and congestion on link criticality rankings in transportation networks". In: *Transportation Research Part D: Transport and Environment* 87, p. 102529.

Anda C., Erath A., and Fourie P. J. (2017). "Transport modelling in the age of big data". In: *International Journal of Urban Sciences* 21.sup1, pp. 19–42.

Armijo L. (1966). "Minimization of functions having Lipschitz continuous first partial derivatives". In: *Pacific Journal of mathematics* 16.1, pp. 1–3.

Aslani M., Mesgari M. S., and Wiering M. (2017). "Adaptive traffic signal control with actor-critic methods in a real-world traffic network with different traffic disruption events". In: *Transportation Research Part C: Emerging Technologies* 85, pp. 732–752.

Aw A. and Rascle M. (2000). "Resurrection of" second order" models of traffic flow". In: *SIAM journal on applied mathematics* 60.3, pp. 916–938.

Aydin N. Y., Duzgun H. S., Wenzel F., and Heinimann H. R. (2018). "Integration of stress testing with graph theory to assess the resilience of urban road networks under seismic hazards". In: *Natural Hazards* 91.1, pp. 37–68.

Azad N., Hassini E., and Verma M. (2016). "Disruption risk management in railroad networks: An optimization-based methodology and a case study". In: *Transportation Research Part B: Methodological* 85, pp. 70–88.

Babazadeh A., Javani B., Gentile G., and Florian M. (2020). "Reduced gradient algorithm for user equilibrium traffic assignment problem". In: *Transportmetrica A: Transport Science* 16.3, pp. 1111–1135.

Bar-Gera H and Boyce D (1999). "Route flow entropy maximization in origin-based traffic assignment". In: *14th International Symposium on Transportation and Traffic TheoryTransportation Research Institute*.

Bar-Gera H. (2002). "Origin-based algorithm for the traffic assignment problem". In: *Transportation Science* 36.4, pp. 398–417.

— (2005). "Continuous and discrete trajectory models for dynamic traffic assignment". In: *Networks and Spatial Economics* 5.1, pp. 41–70.

— (2006). "Primal method for determining the most likely route flows in large road networks". In: *Transportation Science* 40.3, pp. 269–286.

— (2010). “Traffic assignment by paired alternative segments”. In: *Transportation Research Part B: Methodological* 44.8-9, pp. 1022–1046.

Baroud H., Barker K., Ramirez-Marquez J. E., et al. (2014). “Importance measures for inland waterway network resilience”. In: *Transportation research part E: logistics and transportation review* 62, pp. 55–67.

Beckmann M., McGuire C. B., and Winsten C. B. (1956). *Studies in the Economics of Transportation*. Tech. rep.

Beiler M. O., McNeil S., Ames D., and Gayley R. (2013). “Identifying resiliency performance measures for megaregional planning: Case study of the transportation corridor between Boston, Massachusetts, and Washington, DC”. In: *Transportation research record* 2397.1, pp. 153–160.

Bekhor S and Prashker J. (2001). “Stochastic user equilibrium formulation for generalized nested logit model”. In: *Transportation Research Record* 1752.1, pp. 84–90.

Bekhor S., Ben-Akiva M. E., and Ramming M. S. (2006). “Evaluation of choice set generation algorithms for route choice models”. In: *Annals of Operations Research* 144.1, pp. 235–247.

Bekhor S., Reznikova L., and Toledo T. (2007). “Application of cross-nested logit route choice model in stochastic user equilibrium traffic assignment”. In: *Transportation Research Record* 2003.1, pp. 41–49.

Bekhor S., Toledo T., and Prashker J. N. (2008). “Effects of choice set size and route choice models on path-based traffic assignment”. In: *Transportmetrica* 4.2, pp. 117–133.

Bell M. G. (1999). “Measuring network reliability: a game theoretic approach”. In: *Journal of advanced transportation* 33.2, pp. 135–146.

Bell M. G. and Cassir C. (2002). “Risk-averse user equilibrium traffic assignment: an application of game theory”. In: *Transportation Research Part B: Methodological* 36.8, pp. 671–681.

Bell M. G., Kurauchi F., Perera S., and Wong W. (2017). “Investigating transport network vulnerability by capacity weighted spectral analysis”. In: *Transportation Research Part B: Methodological* 99, pp. 251–266.

Bellman R. (1958). “On a routing problem”. In: *Quarterly of applied mathematics* 16.1, pp. 87–90.

Ben-Akiva M. and Bierlaire M. (1999). “Discrete choice methods and their applications to short term travel decisions”. In: *Handbook of transportation science*. Springer, pp. 5–33.

— (2003). “Discrete choice models with applications to departure time and route choice”. In: *Handbook of transportation science*. Springer, pp. 7–37.

Ben-Akiva M., Bierlaire M., Koutsopoulos H., and Mishalani R. (1998). “DynaMIT: a simulation-based system for traffic prediction”. In: *DACCORD short term forecasting workshop*. Vol. 12. Delft The Netherlands.

Ben-Akiva M., De Palma A., and Isam K. (1991). “Dynamic network models and driver information systems”. In: *Transportation Research Part A: General* 25.5, pp. 251–266.

Ben-Akiva M. and Steven R. L. (1985). *Discrete choice analysis: theory and application to travel demand*. Vol. 9. MIT press.

Ben-Elia E. and Shiftan Y. (2010). “Which road do I take? A learning-based model of route-choice behavior with real-time information”. In: *Transportation Research Part A: Policy and Practice* 44.4, pp. 249–264.

Berche B., Von Ferber C., Holovatch T., and Holovatch Y. (2009). “Resilience of public transport networks against attacks”. In: *The European Physical Journal B* 71.1, pp. 125–137.

Berdica K. (2002). “An introduction to road vulnerability: what has been done, is done and should be done”. In: *Transport policy* 9.2, pp. 117–127.

Bešinović N. (2020). “Resilience in railway transport systems: a literature review and research agenda”. In: *Transport Reviews* 40.4, pp. 457–478.

Bhavathrathan B. and Patil G. R. (2015). “Quantifying resilience using a unique critical cost on road networks subject to recurring capacity disruptions”. In: *Transportmetrica A: Transport Science* 11.9, pp. 836–855.

Bifulco G. N., Di Pace R., and Viti F. (2014). “Evaluating the effects of information reliability on travellers’ route choice”. In: *European Transport Research Review* 6.1, pp. 61–70.

Bingfeng S., Ming Z., Xiaobao Y., and Ziyou G. (2017). “Bi-level programming model for exclusive bus lanes configuration in multimodal traffic network”. In: *Transportation research procedia* 25, pp. 652–663.

Birge J. R. and Ho J. K. (1993). “Optimal flows in stochastic dynamic networks with congestion”. In: *Operations Research* 41.1, pp. 203–216.

Black J. (2018). *Urban transport planning: Theory and practice*. Vol. 4. Routledge.

Bliemer M. C., Raadsen M. P., Brederode L. J., Bell M. G., Wismans L. J., and Smith M. J. (2017). "Genetics of traffic assignment models for strategic transport planning". In: *Transport reviews* 37.1, pp. 56–78.

Bocchini P. and Frangopol D. M. (2012a). "Optimal resilience-and cost-based postdisaster intervention prioritization for bridges along a highway segment". In: *Journal of Bridge Engineering* 17.1, pp. 117–129.

— (2012b). "Restoration of bridge networks after an earthquake: Multicriteria intervention optimization". In: *Earthquake Spectra* 28.2, pp. 427–455.

Bogers E. A., Viti F., and Hoogendoorn S. P. (2005). "Joint modeling of advanced travel information service, habit, and learning impacts on route choice by laboratory simulator experiments". In: *Transportation Research Record* 1926.1, pp. 189–197.

Borchers M., Breeuwsma P., Kern W., Slootbeek J., Still G., and Tibben W. (2015). "Traffic user equilibrium and proportionality". In: *Transportation Research Part B: Methodological* 79, pp. 149–160.

Bovy P. H. and Stern E. (2012). *Route choice: Wayfinding in transport networks: Wayfinding in transport networks*. Vol. 9. Springer Science & Business Media.

Bovy P. H. L., Bekhor S., and Prato C. G. (2008). "The Factor of Revisited Path Size: Alternative Derivation". In: *Transportation Research Record* 2076.1, pp. 132–140. DOI: 10.3141/2076-15.

Bovy P. H. (2009). "On modelling route choice sets in transportation networks: a synthesis". In: *Transport reviews* 29.1, pp. 43–68.

Boyce D. E. (1984). "Urban transportation network-equilibrium and design models: recent achievements and future prospects". In: *Environment and Planning A* 16.11, pp. 1445–1474.

Branston D. (1976). "Models of single lane time headway distributions". In: *Transportation Science* 10.2, pp. 125–148.

Bruneau M., Chang S. E., Eguchi R. T., Lee G. C., O'Rourke T. D., Reinhorn A. M., Shinozuka M., Tierney K., Wallace W. A., and Von Winterfeldt D.

(2003). “A framework to quantitatively assess and enhance the seismic resilience of communities”. In: *Earthquake spectra* 19.4, pp. 733–752.

Bucchiarone A., De Sanctis M., and Bencomo N. (2020). “Agent-Based Framework for Self-Organization of Collective and Autonomous Shuttle Fleets”. In: *IEEE Transactions on Intelligent Transportation Systems* 22.6, pp. 3631–3643.

Burrell J. E. (1968). “Multiple route assignment and its application to capacity restraint”. In: *Proceedings of Fourth International Symposium on the Theory of Traffic Flow*.

Caballini C., Fioribello S., Sacone S., and Siri S. (2016). “An MILP optimization problem for sizing port rail networks and planning shunting operations in container terminals”. In: *IEEE Transactions on Automation Science and Engineering* 13.4, pp. 1492–1503.

Caballini C., Pasquale C., Sacone S., and Siri S. (2013). “An event-triggered receding-horizon scheme for planning rail operations in maritime terminals”. In: *IEEE Transactions on Intelligent Transportation Systems* 15.1, pp. 365–375.

Calvert S. C. and Snelder M. (2018). “A methodology for road traffic resilience analysis and review of related concepts”. In: *Transportmetrica A: transport science* 14.1-2, pp. 130–154.

Cantarella G. E. and Cascetta E. (1995). “Dynamic processes and equilibrium in transportation networks: towards a unifying theory”. In: *Transportation Science* 29.4, pp. 305–329.

Cantarella G. E. and Watling D. P. (2016). “Modelling road traffic assignment as a day-to-day dynamic, deterministic process: a unified approach to discrete-and continuous-time models”. In: *EURO Journal on Transportation and Logistics* 5.1, pp. 69–98.

Cantillo V., Ortuzar J. d. D., and Williams H. C. (2007). “Modeling discrete choices in the presence of inertia and serial correlation”. In: *Transportation Science* 41.2, pp. 195–205.

Carrion C. and Levinson D. M. (2012). “Route choice dynamics after a link restoration”. In.

Carroll Jr J. D. (1959). “A method of traffic assignment to an urban network”. In: *Highway Research Board Bulletin* 224.

Casas J., Ferrer J. L., Garcia D., Perarnau J., and Torday A. (2010). "Traffic simulation with aimsun". In: *Fundamentals of traffic simulation*. Springer, pp. 173–232.

Cascetta E. (1989). "A stochastic process approach to the analysis of temporal dynamics in transportation networks". In: *Transportation Research Part B: Methodological* 23.1, pp. 1–17.

— (2009). *Transportation systems analysis: models and applications*. Vol. 29. Springer Science & Business Media.

— (2013). *Transportation systems engineering: theory and methods*. Vol. 49. Springer Science & Business Media.

Cascetta E., Nuzzolo A., Russo F., and Vitetta A. (1996). "A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks". In: *Transportation and Traffic Theory. Proceedings of The 13th International Symposium On Transportation And Traffic Theory, Lyon, France, 24-26 July 1996*.

Cascetta E. and Papola A. (2001). "Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand". In: *Transportation Research Part C: Emerging Technologies* 9.4, pp. 249–263.

Castillo E., Menéndez J. M., Jiménez P., and Rivas A. (2008). "Closed form expressions for choice probabilities in the Weibull case". In: *Transportation Research Part B: Methodological* 42.4, pp. 373–380.

Chaniotakis E. and Pel A. J. (2015). "Drivers' parking location choice under uncertain parking availability and search times: A stated preference experiment". In: *Transportation Research Part A: Policy and Practice* 82, pp. 228–239.

Chen A., Jayakrishnan R, and Tsai W. K. (2002). "Faster Frank-Wolfe traffic assignment with new flow update scheme". In: *Journal of Transportation Engineering* 128.1, pp. 31–39.

Chen A., Xu X., Ryu S., and Zhou Z. (2013). "A self-adaptive Armijo stepsize strategy with application to traffic assignment models and algorithms". In: *Transportmetrica A: Transport Science* 9.8, pp. 695–712.

Chen A., Yang C., Kongsomsaksakul S., and Lee M. (2007). "Network-based accessibility measures for vulnerability analysis of degradable transportation networks". In: *Networks and Spatial Economics* 7.3, pp. 241–256.

Chen B. Y., Lam W. H., Sumalee A., Li Q., and Li Z.-C. (2012). "Vulnerability analysis for large-scale and congested road networks with demand uncertainty". In: *Transportation Research Part A: Policy and Practice* 46.3, pp. 501–516.

Chen P., Scown C., Matthews H. S., Garrett Jr J. H., and Hendrickson C. (2009). "Managing critical infrastructure interdependence through economic input-output methods". In: *Journal of Infrastructure Systems* 15.3, pp. 200–210.

Cho H.-J. and Hwang M.-C. (2005). "A stimulus-response model of day-to-day network dynamics". In: *IEEE Transactions on Intelligent Transportation Systems* 6.1, pp. 17–25.

Chopra S. S., Dillon T., Bilec M. M., and Khanna V. (2016). "A network-based framework for assessing infrastructure resilience: a case study of the London metro system". In: *Journal of The Royal Society Interface* 13.118, p. 20160113.

Chunyan L., Jun C., Zheng'an S., and Xiaofei Y. (2016). "Drivers' route choice behavior analysis under ATIS". In: *Journal of Shenzhen University Science and Engineering [ISSN: 1000-2618/CN: 44-1401/N]* 33.2, pp. 111–220.

Clark C. E. (1961). "The greatest of a finite set of random variables". In: *Operations Research* 9.2, pp. 145–162.

Clark S. D. and Watling D. P. (2002). "Sensitivity analysis of the probit-based stochastic user equilibrium assignment model". In: *Transportation Research Part B: Methodological* 36.7, pp. 617–635.

Conlisk J. (1996). "Why bounded rationality?" In: *Journal of economic literature* 34.2, pp. 669–700.

Corthout R., Flötteröd G., Viti F., and Tampère C. M. (2012). "Non-unique flows in macroscopic first-order intersection models". In: *Transportation Research Part B: Methodological* 46.3, pp. 343–359.

Dafermos S. (1980). "Traffic equilibrium and variational inequalities". In: *Transportation science* 14.1, pp. 42–54.

— (1982a). "Relaxation algorithms for the general asymmetric traffic equilibrium problem". In: *Transportation Science* 16.2, pp. 231–240.

— (1982b). "The general multimodal network equilibrium problem with elastic demand". In: *Networks* 12.1, pp. 57–72.

Dafermos S. C. (1976). "Integrated equilibrium flow models for transportation planning". In: *Traffic Equilibrium Methods*. Springer, pp. 106–118.

Dafermos S. C. and Sparrow F. T. (1969). "The traffic assignment problem for a general network". In: *Journal of Research of the National Bureau of Standards B* 73.2, pp. 91–118.

Daganzo C. F. (1977). "Some statistical problems in connection with traffic assignment". In: *Transportation Research* 11.6, pp. 385–389.

— (1994). "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory". In: *Transportation Research Part B: Methodological* 28.4, pp. 269–287.

Daganzo C. F. and Sheffi Y. (1977). "On stochastic models of traffic assignment". In: *Transportation science* 11.3, pp. 253–274.

Davis G. A. and Nihan N. L. (1993). "Large population approximations of a general stochastic traffic assignment model". In: *Operations Research* 41.1, pp. 169–178.

De La Barra T., Perez B, and Anez J (1993). "Multidimensional path search and assignment". In: *PTRC Summer Annual Meeting, 21st, 1993, University of Manchester, United Kingdom*.

De Palma A. and Lindsey R. (2011). "Traffic congestion pricing methodologies and technologies". In: *Transportation Research Part C: Emerging Technologies* 19.6, pp. 1377–1399.

Demšar U., Špatenková O., and Virrantaus K. (2008). "Identifying critical locations in a spatial network with graph theory". In: *Transactions in GIS* 12.1, pp. 61–82.

Di X., He X., Guo X., and Liu H. X. (2014). "Braess paradox under the boundedly rational user equilibria". In: *Transportation Research Part B: Methodological* 67, pp. 86–108.

Di X. and Liu H. X. (2016). "Boundedly rational route choice behavior: A review of models and methodologies". In: *Transportation Research Part B: Methodological* 85, pp. 142–179.

Di X., Liu H. X., Ban X. J. X., and Yu J. W. (2015). "Submission to the DTA 2012 special issue: On the stability of a boundedly rational day-to-day dynamic". In: *Networks and Spatial Economics* 15.3, pp. 537–557.

Dial R. B. (1971). "A probabilistic multipath traffic assignment model which obviates path enumeration". In: *Transportation research* 5.2, pp. 83–111.

— (1997). "Bicriterion traffic assignment: efficient algorithms plus examples". In: *Transportation Research Part B: Methodological* 31.5, pp. 357–379.

— (2006). "A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration". In: *Transportation Research Part B: Methodological* 40.10, pp. 917–936.

Dijkstra E. W. et al. (1959). "A note on two problems in connexion with graphs". In: *Numerische mathematik* 1.1, pp. 269–271.

Dorbritz R. (2011). "Assessing the resilience of transportation systems in case of large-scale disastrous events". In: *Environmental Engineering. Proceedings of the International Conference on Environmental Engineering. ICEE*. Vol. 8. Vilnius Gediminas Technical University, Department of Construction Economics ..., p. 1070.

Duan Y. and Lu F. (2014). "Robustness of city road networks at different granularities". In: *Physica A: Statistical Mechanics and its Applications* 411, pp. 21–34.

D'Lima M. and Medda F. (2015). "A new measure of resilience: An application to the London Underground". In: *Transportation Research Part A: Policy and Practice* 81, pp. 35–46.

Eash R, Chon K, Lee Y, and Boyce D (1979). "Equilibrium traffic assignment on an aggregated highway network for sketch planning". In: *Transportation Research* 13, pp. 243–257.

Elbery A., Dvorak F., Du J., Rakha H. A., and Klenk M. (2018). "Large-scale Agent-based Multi-modal Modeling of Transportation Networks-System Model and Preliminary Results." In: *VEHITS*, pp. 103–112.

Erath A., Birdsall J., Axhausen K. W., and Hajdin R. (2009). "Vulnerability assessment methodology for Swiss road network". In: *Transportation Research Record* 2137.1, pp. 118–126.

Evans L. and Rothery R. (1973). *Experimental measurements of perceptual thresholds in car-following*. 464.

Falkenhausen H. von (1968). *Traffic assignment by a stochastic model*.

Faturechi R. and Miller-Hooks E. (2014). "Travel time resilience of roadway networks under disaster". In: *Transportation research part B: methodological* 70, pp. 47–64.

Fellendorf M. and Vortisch P. (2010). "Microscopic traffic flow simulator VISSIM". In: *Fundamentals of traffic simulation*. Springer, pp. 63–93.

Ferrara A., Sacone S., and Siri S. (2018). *Freeway traffic modelling and control*. Springer.

Ferro G., Minciardi R., Parodi L., and Robba M. (2021). "Optimal Planning of Charging Stations in Coupled Transportation and Power Networks Based on User Equilibrium Conditions". In: *IEEE Transactions on Automation Science and Engineering*.

Fisk C. (1980). "Some developments in equilibrium traffic assignment". In: *Transportation Research Part B: Methodological* 14.3, pp. 243–255.

Fisk C. and Nguyen S. (1982). "Solution algorithms for network equilibrium models with asymmetric user costs". In: *Transportation Science* 16.3, pp. 361–381.

Fisk C. and Boyce D. E. (1983). "Alternative variational inequality formulations of the network equilibrium-travel choice problem". In: *Transportation Science* 17.4, pp. 454–463.

Florian M. (1977). "A traffic equilibrium model of travel by car and public transit modes". In: *Transportation Science* 11.2, pp. 166–179.

Florian M., Constantin I., and Florian D. (2009). "A new look at projected gradient method for equilibrium assignment". In: *Transportation Research Record* 2090.1, pp. 10–16.

Florian M. and Fox B. (1976). "On the probabilistic origin of Dial's multipath traffic assignment model". In: *Transportation Research* 10.5, pp. 339–341.

Florian M., Guálat J., and Spiess H. (1987). "An efficient implementation of the "Partan" variant of the linear approximation method for the network equilibrium problem". In: *Networks* 17.3, pp. 319–339.

Florian M. and Nguyen S. (1978). "A combined trip distribution modal split and trip assignment model". In: *Transportation Research* 12.4, pp. 241–246.

Fonzone A. and Bell M. G. (2010). "Bounded rationality in hyperpath assignment: the locally rational traveller model". In: *Proceedings of TRB 89th Annual Meeting, Washington, DC*.

Frank M. and Wolfe P. (1956). "An algorithm for quadratic programming". In: *Naval research logistics quarterly* 3.1-2, pp. 95–110.

Friesz T. L., Bernstein D., Mehta N. J., Tobin R. L., and Ganjalizadeh S. (1994). "Day-to-day dynamic network disequilibria and idealized traveler information systems". In: *Operations Research* 42.6, pp. 1120–1136.

Friesz T. L., Bernstein D., Smith T. E., Tobin R. L., and Wie B.-W. (1993). "A variational inequality formulation of the dynamic network user equilibrium problem". In: *Operations research* 41.1, pp. 179–191.

Friesz T. L., Han K., Neto P. A., Meimand A., and Yao T. (2013). "Dynamic user equilibrium based on a hydrodynamic model". In: *Transportation Research Part B: Methodological* 47, pp. 102–126.

Friesz T. L., Kim T., Kwon C., and Rigdon M. A. (2011). "Approximate network loading and dual-time-scale dynamic user equilibrium". In: *Transportation Research Part B: Methodological* 45.1, pp. 176–207.

Friesz T. L., Luque J., Tobin R. L., and Wie B.-W. (1989). "Dynamic network traffic assignment considered as a continuous time optimal control problem". In: *Operations Research* 37.6, pp. 893–901.

Fu X., Lam W. H., and Chen B. Y. (2014). "A reliability-based traffic assignment model for multi-modal transport network under demand uncertainty". In: *Journal of Advanced Transportation* 48.1, pp. 66–85.

Fukushima M. (1984). "A modified Frank-Wolfe algorithm for solving the traffic assignment problem". In: *Transportation Research Part B: Methodological* 18.2, pp. 169–177.

Gabriel S. A. and Bernstein D. (1997). "The traffic equilibrium problem with nonadditive path costs". In: *Transportation Science* 31.4, pp. 337–348.

Galligari A. and Sciandrone M. (2018). "A convergent and fast path equilibration algorithm for the traffic assignment problem". In: *Optimization Methods and Software* 33.2, pp. 354–371.

Ganić E., Babić O., Čangalović M., and Stanojević M. (2018). "Air traffic assignment to reduce population noise exposure using activity-based approach". In: *Transportation Research Part D: Transport and Environment* 63, pp. 58–71.

García-Palomares J. C., Gutiérrez J., Martín J. C., and Moya-Gómez B. (2018). "An analysis of the Spanish high capacity road network criticality". In: *Transportation* 45.4, pp. 1139–1159.

Gauthier P., Furno A., and El Faouzi N.-E. (2018). "Road network resilience: how to identify critical links subject to day-to-day disruptions". In: *Transportation research record* 2672.1, pp. 54–65.

Gentile G. (2014). "Local user cost equilibrium: a bush-based algorithm for traffic assignment". In: *Transportmetrica A: Transport Science* 10.1, pp. 15–54.

Gentile G. et al. (2010). "The general link transmission model for dynamic network loading and a comparison with the DUE algorithm". In: *New developments in transport planning: advances in Dynamic Traffic Assignment* 178, p. 153.

Gibb J. (2011). "Model of traffic flow capacity constraint through nodes for dynamic network loading with queue spillback". In: *Transportation research record* 2263.1, pp. 113–122.

Gifford J. L. and Checherita C. (2007). *Bounded rationality and transportation behavior: lessons for public policy*. Tech. rep.

Gille A., Klemenz M., and Siefer T. (2008). "Applying multiscaling analysis to detect capacity resources in railway networks". In: *WIT Transactions on The Built Environment* 103, pp. 595–603.

Gipps P. G. (1981). "A behavioural car-following model for computer simulation". In: *Transportation Research Part B: Methodological* 15.2, pp. 105–111.

— (1986). "A model for the structure of lane-changing decisions". In: *Transportation Research Part B: Methodological* 20.5, pp. 403–414.

Goldstein A. A. (1965). "On steepest descent". In: *Journal of the Society for Industrial and Applied Mathematics, Series A: Control* 3.1, pp. 147–151.

González Ramírez H. (2020). "Study of the choice behaviour of travellers in a transport network via a "simulation game"". PhD thesis. Lyon.

Gunnarson S (1972). "An algorithm for multipath traffic assignment". In: *PTRC Seminar Proceedings, Urban Traffic Model Research*.

Guo R.-Y. and Huang H.-J. (2009). "Chaos and bifurcation in dynamical evolution process of traffic assignment with flow "mutation"". In: *Chaos, Solitons & Fractals* 41.3, pp. 1150–1157.

— (2016). "A discrete dynamical system of formulating traffic assignment: Revisiting Smith's model". In: *Transportation Research Part C: Emerging Technologies* 71, pp. 122–142.

Guo R.-Y., Yang H., and Huang H.-J. (2013). "A discrete rational adjustment process of link flows in traffic networks". In: *Transportation Research Part C: Emerging Technologies* 34, pp. 121–137.

Guo R.-Y., Yang H., Huang H.-J., and Tan Z. (2015). "Link-based day-to-day network traffic dynamics and equilibria". In: *Transportation Research Part B: Methodological* 71, pp. 248–260.

Guo X. (2013). "Toll sequence operation to realize target flow pattern under bounded rationality". In: *Transportation Research Part B: Methodological* 56, pp. 203–216.

Guo X. and Liu H. X. (2011a). "Bounded rationality and irreversible network change". In: *Transportation Research Part B: Methodological* 45.10, pp. 1606–1618.

— (2011b). "Day-to-Day Dynamic Model in Discrete–Continuum Transportation Networks". In: *Transportation research record* 2263.1, pp. 66–72.

Hadjidimitriou S. N., Dell'Amico M., Cantelmo G., and Viti F. (2015). "Assessing the consistency between observed and modelled route choices through GPS data". In: *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. IEEE, pp. 216–222.

Hall R. W. (1993). "Travel outcome and performance: the effect of uncertainty on accessibility". In: *Transportation Research Part B: Methodological* 17.4, pp. 275–290.

Hamdar S. H. and Mahmassani H. S. (2008). "From existing accident-free car-following models to colliding vehicles: exploration and assessment". In: *Transportation research record* 2088.1, pp. 45–56.

Hammond J. H. (1984). "Solving asymmetric variational inequality problems and systems of equations with generalized nonlinear programming algorithms". PhD thesis. Massachusetts Institute of Technology.

Han L. and Du L. (2012). “On a link-based day-to-day traffic assignment model”. In: *Transportation Research Part B: Methodological* 46.1, pp. 72–84.

Hasan S. and Foliente G. (2015). “Modeling infrastructure system interdependencies and socioeconomic impacts of failure in extreme events: emerging R&D challenges”. In: *Natural Hazards* 78.3, pp. 2143–2168.

Hazelton M. L. (2022). “The emergence of stochastic user equilibria in day-to-day traffic models”. In: *Transportation Research Part B: Methodological* 158, pp. 102–112.

Hazelton M. L. and Watling D. P. (2004). “Computation of equilibrium distributions of Markov traffic-assignment models”. In: *Transportation Science* 38.3, pp. 331–342.

HCM (2000). “Highway Capacity Manual”. In: *Washington, DC* 2, p. 1.

He X., Guo X., and Liu H. X. (2010). “A link-based day-to-day traffic assignment model”. In: *Transportation Research Part B: Methodological* 44.4, pp. 597–608.

He X. and Liu H. X. (2012). “Modeling the day-to-day traffic evolution process after an unexpected network disruption”. In: *Transportation Research Part B: Methodological* 46.1, pp. 50–71.

He X. and Peeta S. (2016). “A marginal utility day-to-day traffic evolution model based on one-step strategic thinking”. In: *Transportation Research Part B: Methodological* 84, pp. 237–255.

Hearn D., Lawphongpanich S, and Ventura J. (1985). “Finiteness in restricted simplicial decomposition”. In: *Operations Research Letters* 4.3, pp. 125–130.

Hiraoka T., Iritani I., Okabe K., and Kumamoto H. (2002). “Route choice behavior model based on bounded rationality”. In: *Proc. of 10th ITS World Congress, CD-ROM*.

Hitchcock F. L. (1941). “The distribution of a product from several sources to numerous localities”. In: *Journal of mathematics and physics* 20.1-4, pp. 224–230.

Holling C. S. (1973). “Resilience and stability of ecological systems”. In: *Annual review of ecology and systematics* 4.1, pp. 1–23.

Hoogendoorn S. P. and Bovy P. H. (2001). "State-of-the-art of vehicular traffic flow modelling". In: *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 215.4, pp. 283–303.

Horowitz J. L. (1984). "The stability of stochastic equilibrium in a two-link transportation network". In: *Transportation Research Part B: Methodological* 18.1, pp. 13–28.

Hosseini S., Barker K., and Ramirez-Marquez J. E. (2016). "A review of definitions and measures of system resilience". In: *Reliability Engineering & System Safety* 145, pp. 47–61.

Huang H.-J. and Lam W. H. (2002). "Modeling and solving the dynamic user equilibrium route and departure time choice problem in network with queues". In: *Transportation Research Part B: Methodological* 36.3, pp. 253–273.

Huang Y. and Kockelman K. M. (2020). "Electric vehicle charging station locations: Elastic demand, station congestion, and network equilibrium". In: *Transportation Research Part D: Transport and Environment* 78, p. 102179.

Husdal J. et al. (2004). "Reliability/vulnerability versus cost/benefit". In: *Nicholson AJ, Dante, A*.

Janson B. N. (1991). "Dynamic traffic assignment for urban road networks". In: *Transportation Research Part B: Methodological* 25.2-3, pp. 143–161.

Javani B. and Babazadeh A. (2017). "Origin-destination-based truncated quadratic programming algorithm for traffic assignment problem". In: *Transportation Letters* 9.3, pp. 166–176.

Jayakrishnan R, Tsai W. T., Prashker J. N., and Rajadhyaksha S. (1994). "A faster path-based algorithm for traffic assignment". In.

Jenelius E. (2010). "User inequity implications of road network vulnerability". In: *Journal of Transport and Land Use* 2.3/4, pp. 57–73.

Jenelius E. and Mattsson L.-G. (2012). "Road network vulnerability analysis of area-covering disruptions: A grid-based approach with case study". In: *Transportation research part A: policy and practice* 46.5, pp. 746–760.

— (2015). "Road network vulnerability analysis: Conceptualization, implementation and application". In: *Computers, environment and urban systems* 49, pp. 136–147.

Jenelius E., Petersen T., and Mattsson L.-G. (2006). "Importance and exposure in road network vulnerability analysis". In: *Transportation Research Part A: Policy and Practice* 40.7, pp. 537–560.

Jiang L., Mahmassani H. S., and Zhang K. (2011). "Congestion pricing, heterogeneous users, and travel time reliability: Multicriterion dynamic user equilibrium model and efficient implementation for large-scale networks". In: *Transportation research record* 2254.1, pp. 58–67.

Jin W.-L. (2007). "A dynamical system model of the traffic assignment problem". In: *Transportation Research Part B: Methodological* 41.1, pp. 32–48.

Jong G. C. de and Bliemer M. C. (2015). "On including travel time reliability of road traffic in appraisal". In: *Transportation Research Part A: Policy and Practice* 73, pp. 80–95.

Jou R.-C. (2001). "Modeling the impact of pre-trip information on commuter departure time and route choice". In: *Transportation Research Part B: Methodological* 35.10, pp. 887–902.

Jourquin B. and Limbourg S. (2006). "Equilibrium traffic assignment on large Virtual Networks: Implementation issues and limits for multimodal freight transport". In: *European Journal of Transport and Infrastructure Research* 6.3.

Kaplan S. and Garrick B. J. (1981). "On the quantitative definition of risk". In: *Risk analysis* 1.1, pp. 11–27.

Kato T., Uchida K., Lam W. H., and Sumalee A. (2021). "Estimation of the value of travel time and of travel time reliability for heterogeneous drivers in a road network". In: *Transportation* 48.4, pp. 1639–1670.

Kaviani A., Thompson R. G., and Rajabifard A. (2017). "Improving regional road network resilience by optimised traffic guidance". In: *Transportmetrica A: Transport Science* 13.9, pp. 794–828.

Khisty C. J. and Arslan T. (2005). "Possibilities of steering the transportation planning process in the face of bounded rationality and unbounded uncertainty". In: *Transportation Research Part C: Emerging Technologies* 13.2, pp. 77–92.

Kitthamkesorn S. and Chen A. (2013). "A path-size weibit stochastic user equilibrium model". In: *Procedia-Social and Behavioral Sciences* 80, pp. 608–632.

Knoop V. L., Snelder M., Zuylen H. J. van, and Hoogendoorn S. P. (2012). "Link-level vulnerability indicators for real-world networks". In: *Transportation Research Part A: Policy and Practice* 46.5, pp. 843–854.

Kotsialos A., Papageorgiou M., Diakaki C., Pavlis Y., and Middelham F. (2002). "Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET". In: *IEEE Transactions on intelligent transportation systems* 3.4, pp. 282–292.

Kumar A. and Peeta S. (2010). "Slope-based multipath flow update algorithm for static user equilibrium traffic assignment problem". In: *Transportation research record* 2196.1, pp. 1–10.

— (2014a). "Slope-based path shift propensity algorithm for the static traffic assignment problem". In: *International Journal for Traffic and Transport Engineering* 4.3, pp. 297–319.

— (2014b). "Strategies to Enhance the Performance of Path-Based Static Traffic Assignment Algorithms". In: *Computer-Aided Civil and Infrastructure Engineering* 29.5, pp. 330–341.

— (2015a). "A day-to-day dynamical model for the evolution of path flows under disequilibrium of traffic networks with fixed demand". In: *Transportation Research Part B: Methodological* 80, pp. 235–256.

— (2015b). "Entropy weighted average method for the determination of a single representative path flow solution for the static user equilibrium traffic assignment problem". In: *Transportation Research Part B: Methodological* 71, pp. 213–229.

Kusakabe T., Sharyo T., and Asakura Y. (2012). "Effects of traffic incident information on drivers' route choice behaviour in urban expressway network". In: *Procedia-Social and Behavioral Sciences* 54, pp. 179–188.

Kuwahara M. and Akamatsu T. (2001). "Dynamic user optimal assignment with physical queues for a many-to-many OD pattern". In: *Transportation Research Part B: Methodological* 35.5, pp. 461–479.

Lam W. H. and Huang H.-J. (1995). "Dynamic user optimal traffic assignment model for many to one travel demand". In: *Transportation Research Part B: Methodological* 29.4, pp. 243–259.

Lansdowne Z. F. (1981). "Rail freight traffic assignment". In: *Transportation Research Part A: General* 15.2, pp. 183–190.

Larsson T., Lundgren J. T., Rydergren C., and Patriksson M. (2001). "Most likely traffic equilibrium route flows analysis and computation". In: *Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Models*. Springer, pp. 129–159.

Larsson T. and Patriksson M. (1992). "Simplicial decomposition with disaggregated representation for the traffic assignment problem". In: *Transportation Science* 26.1, pp. 4–17.

— (1995). "An augmented Lagrangean dual algorithm for link capacity side constrained traffic assignment problems". In: *Transportation Research Part B: Methodological* 29.6, pp. 433–455.

Latora V. and Marchiori M. (2005). "Vulnerability and protection of infrastructure networks". In: *Physical Review E* 71.1, p. 015103.

Leonard D., Gower P, and Taylor N. (1989). *CONTRAM: structure of the model*. Tech. rep.

Leurent F. (2005). "Contributions to the logit assignment model". In: *Transportation Research Record* 1493, pp–207.

Leurent F. M. (1997). "Curbing the computational difficulty of the logit equilibrium assignment model". In: *Transportation Research Part B: Methodological* 31.4, pp. 315–326.

Li T., Rong L., and Yan K. (2019). "Vulnerability analysis and critical area identification of public transport system: A case of high-speed rail and air transport coupling system in China". In: *Transportation Research Part A: Policy and Practice* 127, pp. 55–70.

Li Y., Tan Z., and Chen Q. (2012). "Dynamics of a transportation network model with homogeneous and heterogeneous users". In: *Discrete Dynamics in Nature and Society* 2012.

Liao T.-Y., Hu T.-Y., and Ko Y.-N. (2018). "A resilience optimization model for transportation networks under disasters". In: *Natural hazards* 93.1, pp. 469–489.

Lighthill M. J. and Whitham G. B. (1955). "On kinematic waves II. A theory of traffic flow on long crowded roads". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 229.1178, pp. 317–345.

Lin D.-Y. and Chang Y.-T. (2018). “Ship routing and freight assignment problem for liner shipping: Application to the Northern Sea Route planning problem”. In: *Transportation Research Part E: Logistics and Transportation Review* 110, pp. 47–70.

Lin D.-Y., Fang J.-H., and Huang K.-L. (2019). “Passenger assignment and pricing strategy for a passenger railway transportation system”. In: *Transportation Letters* 11.6, pp. 320–331.

Lin D.-Y. and Leong P. W. (2014). “An N-path user equilibrium for transportation networks”. In: *Applied Mathematical Modelling* 38.2, pp. 667–682.

Liu H., Han J., Fu Y., Zhou J., Lu X., and Xiong H. (2020). “Multi-modal transportation recommendation with unified route representation learning”. In: *Proceedings of the VLDB Endowment* 14.3, pp. 342–350.

Liu R. and Smith M. (2015). “Route choice and traffic signal control: A study of the stability and instability of a new dynamical model of route choice and traffic signal control”. In: *Transportation Research Part B: Methodological* 77, pp. 123–145.

Liu W. and Geroliminis N. (2017). “Doubly dynamics for multi-modal networks with park-and-ride and adaptive pricing”. In: *Transportation Research Part B: Methodological* 102, pp. 162–179.

Lo H. K., Luo X., and Siu B. W. (2006). “Degradable transport network: travel time budget of travelers with heterogeneous risk aversion”. In: *Transportation Research Part B: Methodological* 40.9, pp. 792–806.

Lo H. K., Yip C., and Wan K. (2003). “Modeling transfer and non-linear fare structure in multi-modal network”. In: *Transportation Research Part B: Methodological* 37.2, pp. 149–170.

Lo H. K. and Szeto W. Y. (2002). “A cell-based dynamic traffic assignment model: formulation and properties”. In: *Mathematical and computer modelling* 35.7-8, pp. 849–865.

Loder A., Ambühl L., Menendez M., and Axhausen K. W. (2017). “Empirics of multi-modal traffic networks–Using the 3D macroscopic fundamental diagram”. In: *Transportation Research Part C: Emerging Technologies* 82, pp. 88–101.

Long J., Huang H.-J., Gao Z., and Szeto W. Y. (2013). "An intersection-movement-based dynamic user optimal route choice problem". In: *Operations Research* 61.5, pp. 1134–1147.

Lotan T. (1997). "Effects of familiarity on route choice behavior in the presence of information". In: *Transportation Research Part C: Emerging Technologies* 5.3-4, pp. 225–243.

Lou Y., Yin Y., and Lawphongpanich S. (2010). "Robust congestion pricing under boundedly rational user equilibrium". In: *Transportation Research Part B: Methodological* 44.1, pp. 15–28.

Luathep P., Sumalee A., Ho H., and Kurauchi F. (2011). "Large-scale road network vulnerability analysis: a sensitivity analysis based approach". In: *Transportation* 38.5, pp. 799–817.

Ma F, Wu Q., and Wang L (2014). "Impact of multi-source guidance information on path change of drivers". In: *Journal of Transportation Systems Engineering and Information Technology* 14, pp. 61–66.

Maher M. (1992). "SAM-A stochastic assignment model". In: *Mathematics in Transport Planning and Control (Institute of Mathematics & Its Applications Conference Series 38)*.

Mahmassani H. and Herman R. (1984). "Dynamic user equilibrium departure time and route choice on idealized traffic arterials". In: *Transportation Science* 18.4, pp. 362–384.

Mahmassani H. S. and Chang G.-L. (1987). "On boundedly rational user equilibrium in transportation systems". In: *Transportation science* 21.2, pp. 89–99.

Mahmassani H. S. and Liu Y.-H. (1999). "Dynamics of commuting decision behaviour under advanced traveller information systems". In: *Transportation Research Part C: Emerging Technologies* 7.2-3, pp. 91–107.

Mahnke R., Kaupužs J., and Lubashevsky I (2005). "Probabilistic description of traffic flow". In: *Physics Reports* 408.1-2, pp. 1–130.

Marcotte P. and Guélat J. (1988). "Adaptation of a modified Newton method for solving the asymmetric traffic equilibrium problem". In: *Transportation Science* 22.2, pp. 112–124.

Marsden G., Frick K. T., May A. D., and Deakin E. (2012). "Bounded rationality in policy learning amongst cities: lessons from the transport sector". In: *Environment and Planning A* 44.4, pp. 905–920.

Mason P. (1972). "Multiflow Multiple Routing Procedure." In: *Traffic Engineering & Control* 14, pp. 280–282.

Mather M., Shafir E., and Johnson M. K. (2003). "Remembering chosen and assigned options". In: *Memory & Cognition* 31.3, pp. 422–433.

Mattsson L.-G. and Jenelius E. (2015). "Vulnerability and resilience of transport systems–A discussion of recent research". In: *Transportation research part A: policy and practice* 81, pp. 16–34.

Mayberry J. P. (1973). "Structural requirements for abstract-mode models of passenger transportation". In: *The demand for travel: Theory and Measurement, DC Heath and Co., Lexington, Mass*.

Merchant D. K. and Nemhauser G. L. (1978). "A model and an algorithm for the dynamic traffic assignment problems". In: *Transportation science* 12.3, pp. 183–199.

Mirchandani P. and Soroush H. (1987). "Generalized traffic equilibrium with probabilistic travel times and perceptions". In: *Transportation Science* 21.3, pp. 133–152.

Mirosavljević P., Gvozdenović S., and Čokorilo O. (2011). "A model of air traffic assignment as part of airport air pollution management system". In: *Aviation* 15.4, pp. 92–100.

Mojtahedi M., Newton S., and Von Meding J. (2017). "Predicting the resilience of transport infrastructure to a natural disaster using Cox's proportional hazards regression model". In: *Natural Hazards* 85.2, pp. 1119–1133.

Moore E. F. (1959). "The shortest path through a maze". In: *Proc. Int. Symp. Switching Theory, 1959*, pp. 285–292.

Mounce R. (2006). "Convergence in a continuous dynamic queueing model for traffic networks". In: *Transportation Research Part B: Methodological* 40.9, pp. 779–791.

Mounce R. and Carey M. (2011). "Route swapping in dynamic traffic networks". In: *Transportation Research Part B: Methodological* 45.1, pp. 102–111.

Munby D. L. (1968). *Transport: selected readings*. Vol. 58. Penguin.

Murray-Tuite P. M. (2006). “A comparison of transportation network resilience under simulated system optimum and user equilibrium conditions”. In: *Proceedings of the 2006 winter simulation conference*. IEEE, pp. 1398–1405.

Nagel K. (1998). “From particle hopping models to traffic flow theory”. In: *Transportation Research Record* 1644.1, pp. 1–9.

Nagel K. and Schreckenberg M. (1992). “A cellular automaton model for freeway traffic”. In: *Journal de physique I* 2.12, pp. 2221–2229.

Nagurney A. and Zhang D. (1997). “Projected dynamical systems in the formulation, stability analysis, and computation of fixed-demand traffic network equilibria”. In: *Transportation Science* 31.2, pp. 147–158.

Nagurney A. B. (1984). “Comparative tests of multimodal traffic equilibrium methods”. In: *Transportation Research Part B: Methodological* 18.6, pp. 469–485.

Nair R., Avetisyan H., and Miller-Hooks E. (2010). “Resilience framework for ports and other intermodal components”. In: *Transportation Research Record* 2166.1, pp. 54–65.

Netjasov F. T. (2008). “A model of air traffic assignment as a measure for mitigating noise at airports: the Zurich Airport case”. In: *Transportation Planning and Technology* 31.5, pp. 487–508.

Newell G. F. (1993). “A simplified theory of kinematic waves in highway traffic, part I: General theory”. In: *Transportation Research Part B: Methodological* 27.4, pp. 281–287.

Nguyen S. and Dupuis C. (1984). “An efficient method for computing traffic equilibria in networks with asymmetric transportation costs”. In: *Transportation Science* 18.2, pp. 185–202.

Nicholson A., Schmöcker J.-D., Bell M. G., and Iida Y. (2003). “Assessing transport reliability: malevolence and user knowledge”. In: *The network reliability of transport*. Emerald Group Publishing Limited.

Nie Y., Zhang H., and Lee D.-H. (2004). “Models and algorithms for the traffic assignment problem with link capacity constraints”. In: *Transportation Research Part B: Methodological* 38.4, pp. 285–312.

Nielsen O. A. (2000). "A stochastic transit assignment model considering differences in passengers utility functions". In: *Transportation Research Part B: Methodological* 34.5, pp. 377–402.

Omer M., Mostashari A., Nilchiani R., and Mansouri M. (2012). "A framework for assessing resiliency of maritime transportation systems". In: *Maritime Policy & Management* 39.7, pp. 685–703.

Ordóñez F. and Stier-Moses N. E. (2007). "Robust wardrop equilibrium". In: *International Conference on Network Control and Optimization*. Springer, pp. 247–256.

Osei-Asamoah A. and Lownes N. E. (2014). "Complex network method of evaluating resilience in surface transportation networks". In: *Transportation Research Record* 2467.1, pp. 120–128.

Overgaard K. R. (1967). "Urban transportation planning'traffic estimation". In: *Traffic Quarterly* 21.2.

Papageorgiou M., Blosseville J.-M., and Hadj-Salem H. (1990). "Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: Part I: Modelling". In: *Transportation Research Part A: General* 24.5, pp. 345–359.

Papinski D., Scott D. M., and Doherty S. T. (2009). "Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based GPS". In: *Transportation research part F: traffic psychology and behaviour* 12.4, pp. 347–358.

Pasquale C., Siri E., Sacone S., and Siri S (2021a). "A Discrete-Time Model for Large-Scale Multi-Modal Transport Networks". In: *IFAC-PapersOnLine* 54.2, pp. 7–12.

— (2021b). "A Modeling Framework for Passengers and Freight in Large-Scale Multi-Modal Transport Networks". In: *2021 29th Mediterranean Conference on Control and Automation (MED)*. IEEE, pp. 681–686.

Pasquale C., Siri E., Siri S., and Sacone S. (2022). "Two-Stage Multi-Class Modeling Approach for Multi-Modal Transport Networks". In: *To be submitted for second round of review to IEEE Access*.

Patriksson M. (2015). *The traffic assignment problem: models and methods*. Courier Dover Publications.

Paveri-Fontana S. (1975). "On Boltzmann-like treatments for traffic flow: a critical review of the basic model and an alternative proposal for dilute traffic analysis". In: *Transportation research* 9.4, pp. 225–235.

Payne H. J. (1979). "FREFLO: A macroscopic simulation model of freeway traffic". In: *Transportation Research Record* 722.

Peeta S. and Yang T.-H. (2003). "Stability issues for dynamic traffic assignment". In: *Automatica* 39.1, pp. 21–34.

Peeta S. and Ziliaskopoulos A. K. (2001). "Foundations of dynamic traffic assignment: The past, the present and the future". In: *Networks and spatial economics* 1.3, pp. 233–265.

Pel A. J., Bliemer M. C., and Hoogendoorn S. P. (2009). "Hybrid route choice modeling in dynamic traffic assignment". In: *Transportation Research Record* 2091.1, pp. 100–107.

Pel A. J. and Chaniotakis E. (2017). "Stochastic user equilibrium traffic assignment with equilibrated parking search routes". In: *Transportation Research Part B: Methodological* 101, pp. 123–139.

Pi X., Ma W., and Qian Z. S. (2019). "A general formulation for multi-modal dynamic traffic assignment considering multi-class vehicles, public transit and parking". In: *Transportation Research Part C: Emerging Technologies* 104, pp. 369–389.

Pigou A. C. (1920). *The economics of welfare*. Palgrave Macmillan.

Potts R. B. and Oliver R. M. (1972). *Flows in transportation networks*. New York, Academic Press.

Powell W. B. and Sheffi Y. (1982). "The convergence of equilibrium algorithms with predetermined step sizes". In: *Transportation Science* 16.1, pp. 45–55.

Prashker J. N. and Bekhor S. (1998). "Investigation of stochastic network loading procedures". In: *Transportation Research Record* 1645.1, pp. 94–102.

Prato C. G. and Bekhor S. (2006). "Applying branch-and-bound technique to route choice set generation". In: *Transportation Research Record* 1985.1, pp. 19–28.

Prato C. G., Bekhor S., and Pronello C. (2012). "Latent variables and route choice behavior". In: *Transportation* 39.2, pp. 299–319.

Rakha H. and Tawfik A. (2019). "Dynamic Traffic Routing, Assignment, and Assessment of Traffic Networks". In: *Complex Dynamics of Traffic Management*. Ed. by B. S. Kerner. New York, NY: Springer US, pp. 79–129. ISBN: 978-1-4939-8763-4. DOI: 10.1007/978-1-4939-8763-4_562. URL: https://doi.org/10.1007/978-1-4939-8763-4_562.

Rakha H., Tawfik A. M., and Meyers R. (2009). *Traffic Networks: Dynamic Traffic Routing, Assignment, and Assessment.*

Ramming M. S. (2001). "Network knowledge and route choice". In: *Unpublished Ph. D. Thesis, Massachusetts Institute of Technology*.

Ran B. and Boyce D. E. (1996). "A link-based variational inequality formulation of ideal dynamic user-optimal route choice problem". In: *Transportation Research Part C: Emerging Technologies* 4.1, pp. 1–12.

Ran B., Boyce D. E., and LeBlanc L. J. (1993). "A new class of instantaneous dynamic user-optimal traffic assignment models". In: *Operations Research* 41.1, pp. 192–202.

Reuschel A (1950). "Vehicle movements in a platoon". In: *Oesterreichisches Ingenieur-Archir* 4, pp. 193–215.

Rossi T. F., McNeil S., and Hendrickson C. (1989). "Entropy model for consistent impact-fee assessment". In: *Journal of urban planning and development* 115.2, pp. 51–63.

Ruiter E. R. (1968). *ICES TRANSET 1. TRANSPORTATION NETWORK ANALYSIS. ENGINEERING USERS'MANUAL*. Tech. rep.

Rupi F., Bernardi S., Rossi G., and Danesi A. (2015). "The evaluation of road network vulnerability in mountainous areas: a case study". In: *Networks and Spatial Economics* 15.2, pp. 397–411.

Samuelson W. and Zeckhauser R. (1988). "Status quo bias in decision making". In: *Journal of risk and uncertainty* 1.1, pp. 7–59.

Sandholm W. H. (2001). "Potential games with continuous player sets". In: *Journal of Economic theory* 97.1, pp. 81–108.

Schintler L. A., Kulkarni R., Gorman S., and Stough R. (2007). "Using raster-based GIS and graph theory to analyze complex networks". In: *Networks and Spatial Economics* 7.4, pp. 301–313.

Scott D. M., Novak D. C., Aultman-Hall L., and Guo F. (2006). "Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks". In: *Journal of Transport Geography* 14.3, pp. 215–227.

Shang W., Han K., Ochieng W., and Angeloudis P. (2017). "Agent-based day-to-day traffic network model with information percolation". In: *Transportmetrica A Transport Science* 13.1, pp. 38–66.

Shao H., Lam W. H., Meng Q., and Tam M. L. (2006). "Demand-driven traffic assignment problem based on travel time reliability". In: *Transportation Research Record* 1985.1, pp. 220–230.

Shao H., Lam W. H., and Tam M. L. (2006). "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand". In: *Networks and Spatial Economics* 6.3, pp. 173–204.

Sheffi Y. (1985). *Urban transportation networks*. Vol. 6. Prentice-Hall, Englewood Cliffs, NJ.

Sheffi Y., Hall R., and Daganzo C. (1982). "On the estimation of the multinomial probit model". In: *Transportation Research Part A: General* 16.5-6, pp. 447–456.

Sheffi Y. and Powell W. B. (1982). "An algorithm for the equilibrium assignment problem with random link times". In: *Networks* 12.2, pp. 191–207.

Sheffi Y. and Daganzo C. F. (1978). "Hypernetworks and supply-demand equilibrium obtained with disaggregate demand models". In: *Transportation Research Record* 673, pp. 113–121.

Shengguo G. and Zhong W. (2011). "Modeling passenger flow distribution based on travel time of urban rail transit". In: *Journal of Transportation Systems Engineering and Information Technology* 11.6, pp. 124–130.

Si B., Zhong M., Yang X., and Gao Z. (2012). "Modeling the congestion cost and vehicle emission within multimodal traffic network under the condition of equilibrium". In: *Journal of Systems Science and Systems Engineering* 21.4, pp. 385–402.

Simon H. A. (1957). "Models of man; social and rational." In.

Siri E., Siri S., and Sacone S. (2020a). "A progressive traffic assignment procedure on networks affected by disruptive events". In: *2020 European Control Conference (ECC)*. IEEE, pp. 130–135.

— (2020b). "Network performance evaluation under disruptive events through a progressive traffic assignment model". In: *IFAC-PapersOnLine* 53.2, pp. 15017–15022.

— (2022a). "A topology-based bounded rationality day-to-day traffic assignment model". In: *Accepted for publication on Communication in Transportation Research*.

— (2022b). "Day-to-day discrete-time traffic assignment model for transport networks affected by disruptive events". In: *European Journal of Control*.

Sivak M. (2002). "How common sense fails us on the road: contribution of bounded rationality to the annual worldwide toll of one million traffic fatalities". In: *Transportation research part F: traffic psychology and behaviour* 5.4, pp. 259–269.

Small K. A. (1982). "The scheduling of consumer activities: work trips". In: *The American Economic Review* 72.3, pp. 467–479.

Smith M. J. (1979). "The existence, uniqueness and stability of traffic equilibria". In: *Transportation Research Part B: Methodological* 13.4, pp. 295–304.

— (1984). "The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov". In: *Transportation science* 18.3, pp. 245–252.

Smith M. J. and Watling D. P. (2016). "A route-swapping dynamical system and Lyapunov function for stochastic user equilibrium". In: *Transportation Research Part B: Methodological* 85, pp. 132–141.

Smith M. and Mounce R. (2011). "A splitting rate model of traffic re-routeing and traffic control". In: *Procedia-Social and Behavioral Sciences* 17, pp. 316–340.

Smith M. J. (1983). "The existence and calculation of traffic equilibria". In: *Transportation Research Part B: Methodological* 17.4, pp. 291–303.

Smith M. and Wisten M. (1995). “A continuous day-to-day traffic assignment model and the existence of a continuous dynamic user equilibrium”. In: *Annals of Operations Research* 60.1, pp. 59–79.

Smock R. B. (1963). “A comparative description of a capacity-restrained traffic assignment”. In: *Highway Research Record* 6.

Snelder M, Van Zuylen H., and Immers L. (2012). “A framework for robustness analysis of road networks for short term variations in supply”. In: *Transportation Research Part A: Policy and Practice* 46.5, pp. 828–842.

Srinivasan K. K. and Mahmassani H. S. (1999). “Role of congestion and information in trip-makers’ dynamic decision processes: Experimental investigation”. In: *Transportation Research Record* 1676.1, pp. 44–52.

Stamos I., Mitsakis E., Salanova J. M., and Aifadopoulou G. (2015). “Impact assessment of extreme weather events on transport networks: A data-driven approach”. In: *Transportation research part D: transport and environment* 34, pp. 168–178.

Stouffer S. A. (1940). “Intervening opportunities: a theory relating mobility and distance”. In: *American sociological review* 5.6, pp. 845–867.

Study (1960). “Chicago Area Transportation Study, Final Report, Volume II: Data Prjections”. In: *Illinois Department of Public Works and Buildings*.

Sullivan J., Aultman-Hall L., and Novak D. (2009). “A review of current practice in network disruption analysis and an assessment of the ability to account for isolating links in transportation networks”. In: *Transportation Letters* 1.4, pp. 271–280.

Sullivan J. L., Novak D. C., Aultman-Hall L., and Scott D. M. (2010). “Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach”. In: *Transportation Research Part A: Policy and Practice* 44.5, pp. 323–336.

Szeto W and Wong S (2012). “Dynamic traffic assignment: model classifications and recent advances in travel choice principles”. In: *Open Engineering* 2.1, pp. 1–18.

Szeto W. Y. and Lo H. K. (2004). “A cell-based simultaneous route and departure time choice model with elastic demand”. In: *Transportation Research Part B: Methodological* 38.7, pp. 593–612.

Szeto W. and Lo H. K. (2006). "Dynamic traffic assignment: properties and extensions". In: *Transportmetrica* 2.1, pp. 31–52.

Tampère C. M., Corthout R., Cattrysse D., and Immers L. H. (2011). "A generic class of first order node models for dynamic macroscopic simulation of traffic flows". In: *Transportation Research Part B: Methodological* 45.1, pp. 289–309.

Tan Z., Yang H., and Guo R.-Y. (2015). "Dynamic congestion pricing with day-to-day flow evolution and user heterogeneity". In: *Transportation Research Part C: Emerging Technologies* 61, pp. 87–105.

Tawfik A. M. and Rakha H. A. (2013). "Latent Class choice model of heterogeneous drivers' route choice behavior based on learning in a real-world experiment". In: *Transportation research record* 2334.1, pp. 84–94.

Taylor M. A. (2013). "Travel through time: the story of research on travel time reliability". In: *Transportmetrica B: transport dynamics* 1.3, pp. 174–194.

Taylor M. A. and D'este G. (2003). "Concepts of network vulnerability and applications to the identification of critical elements of transport infrastructure". In: New Zealand Transport Research Forum Wellington, New Zealand.

Taylor M. A., Sekhar S. V., and D'Este G. M. (2006). "Application of accessibility based methods for vulnerability analysis of strategic road networks". In: *Networks and Spatial Economics* 6.3, pp. 267–291.

Taylor M. A. et al. (2012). "Remoteness and accessibility in the vulnerability analysis of regional road networks". In: *Transportation research part A: policy and practice* 46.5, pp. 761–771.

Testa A. C., Furtado M. N., and Alipour A. (2015). "Resilience of coastal transportation networks faced with extreme climatic events". In: *Transportation Research Record* 2532.1, pp. 29–36.

Tobin R. L. (1977). "An extension of Dial's algorithm utilizing a model of tripmakers' perceptions". In: *Transportation Research* 11.5, pp. 337–342.

Tong C. and Wong S. (2000). "A predictive dynamic traffic assignment model in congested capacity-constrained road networks". In: *Transportation Research Part B: Methodological* 34.8, pp. 625–644.

Traffic Assignment M. of (1964). *Bureau of Public Roads, US Department of Commerce, 1964*.

"Traffic assignment: methods, applications, products" (1972). In.

Tu H., Lint H. van, and Zuylen H. van (2008). "The effects of traffic accidents on travel time reliability". In: *2008 11th International IEEE Conference on Intelligent Transportation Systems*. IEEE, pp. 79–84.

Tversky A. and Kahneman D. (1985). "The framing of decisions and the psychology of choice". In: *Behavioral decision making*. Springer, pp. 25–41.

Twumasi-Boakye R. and Sobanjo J. O. (2018). "Resilience of regional transportation networks subjected to hazard-induced bridge damages". In: *Journal of Transportation Engineering, Part A: Systems* 144.10, p. 04018062.

Uchida T. and Iida Y. (1993). "RISK ASSIGNMENT: A NEW TRAFFIC ASSIGNMENT MODEL CONSIDERING THE RISK OF TRAVEL TIME VARIATION." In: *Transportation and Traffic Theory*.

Van Aerde M et al. (2002). *INTEGRATION Release 2.30 for Windows: User's Guide, Vol. I. Fundamental Features*.

Van Vliet D (1973). "Road assignment: A case for incremental loading". In: *GLTS Note 30*. Department of Planning and Transportation, Greater London Council.

Van Vliet D. (1982). "SATURN-a modern assignment model". In: *Traffic Engineering & Control* 23.HS-034 256.

Venturini G., Iris Ç., Kontovas C. A., and Larsen A. (2017). "The multi-port berth allocation problem with speed optimization and emission considerations". In: *Transportation Research Part D: Transport and Environment* 54, pp. 142–159.

Vickrey W. S. (1969). "Congestion theory and transport investment". In: *The American Economic Review* 59.2, pp. 251–260.

Vliet D. v. (1976). "The Choice of Assignment Techniques for Large Networks". In: *Traffic Equilibrium Methods*. Springer, pp. 396–412.

Voorhees A. M. (2013). "A general theory of traffic movement". In: *Transportation* 40.6, pp. 1105–1116.

Vreeswijk J., Thomas T., Berkum E. van, and Arem B. van (2013). "Drivers' perception of route alternatives as indicator for the indifference band". In: *Transportation research record* 2383.1, pp. 10–17.

Wakabayashi H. and Iida Y. (1992). "Upper and lower bounds of terminal reliability of road networks: an efficient method with Boolean algebra". In: *Journal of natural disaster science* 14.1.

Wang D. Z., Liu H., Szeto W., and Chow A. H. (2016). "Identification of critical combination of vulnerable links in transportation networks–a global optimisation approach". In: *Transportmetrica A Transport Science* 12.4, pp. 346–365.

Wang J., Peeta S., and He X. (2019). "Multiclass traffic assignment model for mixed traffic flow of human-driven vehicles and connected and autonomous vehicles". In: *Transportation Research Part B: Methodological* 126, pp. 139–168.

Wang Y., Szeto W. Y., Han K., and Friesz T. L. (2018). "Dynamic traffic assignment: A review of the methodological advances for environmentally sustainable road transportation applications". In: *Transportation Research Part B: Methodological* 111, pp. 370–394.

Wang Y., Liu H., Han K., Friesz T. L., and Yao T. (2015). "Day-to-day congestion pricing and network resilience". In: *Transportmetrica A: Transport Science* 11.9, pp. 873–895.

Wardrop J. G. (1952). "Road paper. some theoretical aspects of road traffic research." In: *Proceedings of the institution of civil engineers* 1.3, pp. 325–362.

Watling D. (1999). "Stability of the stochastic equilibrium assignment problem: a dynamical systems approach". In: *Transportation Research Part B: Methodological* 33.4, pp. 281–312.

Watling D. and Hazelton M. L. (2003). "The dynamics and equilibria of day-to-day assignment models". In: *Networks and Spatial Economics* 3.3, pp. 349–370.

Watling D. P. and Cantarella G. E. (2015). "Model representation & decision-making in an ever-changing world: the role of stochastic process models of transportation systems". In: *Networks and Spatial Economics* 15.3, pp. 843–882.

Watling D. P. and Hazelton M. L. (2018). "Asymptotic approximations of transient behaviour for day-to-day traffic models". In: *Transportation Research Part B: Methodological* 118, pp. 90–105.

Whitham G. B. (1974). *Linear and nonlinear waves*. John Wiley & Sons.

Wieland A. and Wallenburg C. M. (2012). "Dealing with supply chain risks: Linking risk management practices and strategies to performance". In: *International journal of physical distribution & logistics management*.

Wilson A. (1967). "A statistical theory of spatial distribution models". In: *Transportation Research* 1.3, pp. 253–269.

Wu J., Sun H., Wang D. Z., Zhong M., Han L., and Gao Z. (2013). "Bounded-rationality based day-to-day evolution model for travel behavior analysis of urban railway network". In: *Transportation Research Part C: Emerging Technologies* 31, pp. 73–82.

Xu G., Zhao S., Shi F., and Zhang F. (2017). "Cell transmission model of dynamic assignment for urban rail transit networks". In: *Plos one* 12.11, e0188874.

Xu X., Chen A., and Yang C. (2017). "An optimization approach for deriving upper and lower bounds of transportation network vulnerability under simultaneous disruptions of multiple links". In: *Transportation research procedia* 23, pp. 645–663.

Xu X., Qu K., Chen A., and Yang C. (2021). "A new day-to-day dynamic network vulnerability analysis approach with Weibit-based route adjustment process". In: *Transportation Research Part E: Logistics and Transportation Review* 153, p. 102421.

Yamada T., Russ B. F., Castro J., and Taniguchi E. (2009). "Designing multimodal freight transport networks: A heuristic approach and applications". In: *Transportation science* 43.2, pp. 129–143.

Yang F. (2005). *An evolutionary game theory approach to the day-to-day traffic dynamics*. The University of Wisconsin-Madison.

Yang F. and Zhang D. (2009). "Day-to-day stationary link flow pattern". In: *Transportation Research Part B: Methodological* 43.1, pp. 119–126.

Yang Q. and Koutsopoulos H. N. (1996). "A microscopic traffic simulator for evaluation of dynamic traffic management systems". In: *Transportation Research Part C: Emerging Technologies* 4.3, pp. 113–129.

Ye H. (2022). “On Stochastic-User-Equilibrium-Based Day-to-Day Dynamics”. In: *Transportation Science* 56.1, pp. 103–117.

Ye H. and Yang H. (2017). “Rational behavior adjustment process with boundedly rational user equilibrium”. In: *Transportation Science* 51.3, pp. 968–980.

Yildirimoglu M. and Kahraman O. (2018). *Investigating the empirical existence of equilibrium conditions*. Tech. rep.

Yu Y., Han K., and Ochieng W. (2020). “Day-to-day dynamic traffic assignment with imperfect information, bounded rationality and information sharing”. In: *Transportation Research Part C: Emerging Technologies* 114, pp. 59–83.

Zhang D. and Nagurney A. (1996). “On the local and global stability of a travel route choice adjustment process”. In: *Transportation Research Part B: Methodological* 30.4, pp. 245–262.

Zhang D., Nagurney A., and Wu J. (2001). “On the equivalence between stationary link flow patterns and traffic network equilibria”. In: *Transportation Research Part B: Methodological* 35.8, pp. 731–748.

Zhang H. M. (2002). “A non-equilibrium traffic model devoid of gas-like behavior”. In: *Transportation Research Part B: Methodological* 36.3, pp. 275–290.

Zhang J., Liao F., Arentze T., and Timmermans H. (2011). “A multimodal transport network model for advanced traveler information systems”. In: *Procedia-Social and Behavioral Sciences* 20, pp. 313–322.

Zhang J. and Yang H. (2015). “Modeling route choice inertia in network equilibrium with heterogeneous prevailing choice sets”. In: *Transportation Research Part C: Emerging Technologies* 57, pp. 42–54.

Zhang W.-y., Guan W., Ma J.-h., and Tian J.-f. (2015a). “A nonlinear pairwise swapping dynamics to model the selfish rerouting evolutionary game”. In: *Networks and Spatial Economics* 15.4, pp. 1075–1092.

Zhang W., He Z., Guan W., and Qi G. (2018). “Day-to-day rerouting modeling and analysis with absolute and relative bounded rationalities”. In: *Transportmetrica A: Transport Science* 14.3, pp. 256–273.

Zhang W., Jiang S., Ma R., and Guan W. (2019). "A boundedly rational day-to-day rerouting model with dynamic indifference band and net marginal cost preference". In: *Available at SSRN 3275963*.

Zhang X., Mahadevan S., and Goebel K. (2019). "Network reconfiguration for increasing transportation system resilience under extreme events". In: *Risk analysis* 39.9, pp. 2054–2075.

Zhang Y., Yang X., Liu Q., and Li C. (2015b). "Who will use pre-trip traveler information and how will they respond? Insights from Zhongshan metropolitan area, China". In: *Sustainability* 7.5, pp. 5857–5874.

Zheng F., Li J., Van Zuylen H., Liu X., and Yang H. (2018). "Urban travel time reliability at different traffic conditions". In: *Journal of Intelligent Transportation Systems* 22.2, pp. 106–120.

Zheng H. (2015). "Adaptation of network simplex for the traffic assignment problem". In: *Transportation Science* 49.3, pp. 543–558.

Zhou B., Xu M., Meng Q., and Huang Z. (2017). "A day-to-day route flow evolution process towards the mixed equilibria". In: *Transportation Research Part C: Emerging Technologies* 82, pp. 210–228.

Zhou Y., Wang J., and Yang H. (2019). "Resilience of transportation systems: concepts and comprehensive review". In: *IEEE Transactions on Intelligent Transportation Systems* 20.12, pp. 4262–4276.

Zhu L., Yu F. R., Wang Y., Ning B., and Tang T. (2018). "Big data analytics in intelligent transportation systems: A survey". In: *IEEE Transactions on Intelligent Transportation Systems* 20.1, pp. 383–398.

Zhu S. (2010). *The roads taken: Theory and evidence on route choice in the wake of the I-35W Mississippi River bridge collapse and reconstruction*. University of Minnesota.

Zhu S. and Levinson D. (2015). "Do people use the shortest path? An empirical test of Wardrop's first principle". In: *PloS one* 10.8, e0134322.

Zhu Y., Ozbay K., Xie K., and Yang H. (2016). "Using big data to study resilience of taxi and subway trips for hurricanes Sandy and Irene". In: *Transportation research record* 2599.1, pp. 70–80.

Ziliaskopoulos A. and Wardell W. (2000). "An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays". In: *European journal of operational research* 125.3, pp. 486–502.