2020-06-17

Road network vulnerability analysis with consideration of probability and consequences of disruptive events

Ansari Esfeh, Mohammad

UNIVERSITY OF CALGARY

Road network vulnerability analysis with consideration of probability and consequences of disruptive events

by

Mohammad Ansari Esfeh

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CIVIL ENGINEERING

CALGARY, ALBERTA

JUNE, 2020

Abstract

To assess the vulnerability of road networks, the commonly used analytical vulnerability analysis is to determine the global vulnerability of links in a road network by removing links one by one from the network and measuring the resulting increase in the total travel cost of the network. Such global vulnerability ranking might fail to identify the most critical links as it overlooks important factors that affect the vulnerability of road links. Instances of such factors include link specific geometry design, poor downstream traffic signal timing, and/or links that are prone to more collisions. Additionally, traditional techniques identify the critical links of a transportation network measuring only the consequences of the link closure with little consideration given to its closure probability. Consideration of the probability of link closure or failure is important as some of the links in a transportation network are more susceptible to disruptive events than others. To fill the void in the literature, I propose two data-driven vulnerability approaches: 1) vulnerability analysis by modeling monthly and seasonal extreme travel delay variations and 2) vulnerability analysis by measuring the spatiotemporal impact of incidents. In studying road network vulnerability by modeling monthly and seasonal variation of extreme travel delay, I propose a new class of extreme value distribution called compound generalized extreme value (CGEV) distribution for depicting the monthly and seasonal variation in extreme travel delays in road networks. Since the frequency and severity of extreme events are highly correlated to the variation in weather conditions as an extrinsic cause of incidents and long delays, monthly and seasonal changes in weather contribute to extreme travel time variability. The change in driving behavior, which itself varies according to road/weather

conditions, also contributes to the monthly and seasonal variation in observed extreme travel times. Therefore, it is critical to model the effect of monthly and seasonal changes on observed extreme travel delays on road networks. Based on the empirically revealed linear relationship between mean and standard deviation (SD) of extreme travel delays for both monthly and seasonal levels, I formulated two multiplicative error models. I then obtain the CGEV distribution by linking the two multiplicative error models and formed a compound distribution that characterizes the overall variation in extreme travel delay. I calibrated the CGEV distribution parameters and validated the underlying assumptions that are used to derive the CGEV distribution using multi-year observed travel time data from the City of Calgary road network. The results indicate that accounting for the seasonality by identifying seasonal specific parameters provides a flexible and not too complex CGEV distribution that is shown to outperform the traditional GEV distribution. Finally, I evaluated the application of the proposed CGEV distribution in the context of road network vulnerability taking into account the stochastic nature of extreme event occurrences and the link importance. This derived data-driven vulnerability index incorporates a wealth of information related to both network topology in terms of connectivity and the dynamic interaction between travel demand and supply.

In studying road network vulnerability by measuring the spatiotemporal impact of incidents, I propose a new data-driven, impact area, vulnerability analysis approach that takes into consideration both the probability of impact as well as the effects of incidents on the impact area. I employed multi-year observed travel time and incident data to investigate these underlying dynamics as the datasets contain important information that reflect the

historical spatial and temporal occurrences of link closure and their network wide impacts. Rather than focusing solely on the travel time fluctuation of the link subject to incident, and to capture all aspects of incidents' impact, I developed a new approach to identify both the resulting spatial and temporal impacts by monitoring the dynamic propagation and dissemination of congestion patterns in the set of links that are in the vicinity of the link subject to incident (i.e., impact area). I subsequently used these spatial and temporal dimensions of the impacts in the vulnerability analysis. I examined the performance of the developed approach, historical travel time and incident data of the City of Calgary. The results indicate that the recorded temporal impact of incidents is not representative enough of the true impact of incidents since it overlooks the dynamic spatial propagation of the effect of incidents on the impact area. Finally, I used the estimated spatiotemporal impact of incidents in Calgary road network to determine the vulnerability of the links considering both the probability that links in an impact area are affected by an incident as well as the spatiotemporal consequence of the incident on the impact area. This data-driven vulnerability measure could be used as a decision support tool for decision-makers in prioritizing improvements to critical links to enhance overall network vulnerability, reliability, and resilience.

Acknowledgements

I would like to thank my supervisor, Prof. Lina Kattan, for the insightful guidance, encouragement and great support she has provided throughout each stage of my career. I have been extremely lucky to have a supervisor who cared so much about her students. She is not only a passionate and dedicated academic supervisor but an inspiration to me and her students.

I would like to thank Prof. William H.K. Lam for his continuing support. I had the opportunity to work with him while I was in Hong Kong as an exchange student. That exchange program became the starting point for my great research collaboration and I learned a lot from him. My deepest gratitude to Prof. Hp Lo who, although he is no longer with us, continues to inspire by his dedication and contribution to the field of transportation engineering.

I would also like to thank Prof. Chan Wirasinghe, a caring and brilliant scholar. I was very fortunate to work with him on several occasions over the course of my PhD studies at the University of Calgary.

I am very grateful to the members of my Ph.D. defense committee, Drs. Chan Wirasinghe, Nigel Waters, Markus Dann, and Saman Bandara for their valuable advice and helpful suggestions.

I would like to express my gratitude to my friends and teammates at University of Calgary for their constant support and friendship especially Dr. Saeid Saidi, Dr. Mostafa Salari, Shahab Esmaeilnejad and Reza Ansari Esfe, and many friends whom I didn't list here.

My research would not have been practical without valuable data provided by the City of Calgary, INRIX and IBI group to test and validate my model and use the City of Calgary as the case study of this dissertation. I would like to acknowledge Mr. Sameer Patil, the City of Calgary and Mr. Jonathan Darton, IBI group, for providing data and for their support on this research.

Thanks to all current and former staff of the Department of Civil Engineering especially Janelle McConnell, Kate McGillis, Julie Nagy Kovacs, and Chrissy Thatcher who were always ready to help with a smile on their faces.

I am thankful for the support I got from the Natural Science and Engineering Research Council of Canada (NSERC). I also appreciate the support from Mitacs, Eyes High International Doctoral Scholarship, Open Doctoral Scholarship and Cooper H. Langford Graduate Scholarship for their generous support.

This work is financially supported through Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery and Discovery Accelerator Supplement (DAS), Urban Alliance Chair in Transportation Systems Optimization research funds, Mitacs Doctoral Accelerate with IBI group, Calgary and the Alberta Innovate Strategic Research Grant on Integrated Urban Mobility. It is also jointly supported by a PhD student attachment program of the Hong Kong Polytechnic University, together with grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU R5029-18) and from the Research Committee of The Hong Kong Polytechnic University (Project No. 4-ZZFY).

I owe who I am and anything I have achieved to my parents, Hussein Ansari Esfeh and Shahin Fareghi. They are my mentors, my role models, and my inspiration; and I will be in debt to them forever.

Table of Contents

List of Tables

List of Figures and Illustrations

List of Figures and Acronyms

AIC	Akaike Information Criterion
CDF	Cumulative distribution function
CGEV	Compound generalized extreme value
CV	Coefficient of variation
EVT	Extreme value theory
GEV	Generalized extreme value
i.i.d.	Independent and identically distributed
MLE	Maximum likelihood estimate
NMI	normalized mutual information
OD	Origin-destination
PDF	Probability density function
QP	Quadratic optimization problem
q-q	Quantile-quantile
SAIT	Southern Alberta Institute of Technology
SD	Standard deviation

CHAPTER 1: INTRODUCTION

1.1 Background

An urban road network is the backbone of a city and contributes to its operational and economic performance. Yet, the frequent occurrence of various recurrent and non-recurrent events results in severe deterioration to road network performance and its reliability. Recurrent events happen periodically in a road network and are the outcome of the within-day variation of travel demand. Non-recurrent events, in contrast, result from extreme and rare occurrences of events leading to severe deterioration in network performance. The possible sources of such unexpected events are natural (e.g., disasters, extreme weather condition), artificial (e.g., traffic collisions, major road repairs, social events) or technical (e.g., signal failure at a major intersection) (Li, 2008). Other classification of the causes of disruptive events can be further categorised as internal, external and intentional incident interferences (Mattsson and Jenelius, 2015). Internal events are those originating from the system and caused by mistakes made by transportation authorities and staff (e.g., technical failure) and external events are mostly related to natural phenomena (e.g., extreme weather conditions) or artificial events caused by intentional interferences (e.g., terrorist or cyber attacks).

Both recurrent and non-recurrent events can result in long delays and gridlock. Reliability, vulnerability, and resiliency are the three important concepts in the literature that describe the performance of a road network during and after network disruptions. Numerous definitions are proposed in the literature for these three concepts. Reliability is defined as the range of variation of travel time experienced by users during a large number of daily trips (Turner et al., 1996) and indicates the probability that the transportation network remains satisfactory in terms of service

level provision under perturbation (Gu et al., 2019). Vulnerability is defined as the susceptibility of road network elements to perturbation (Taylor, 2017). Finally, resilience refers to the capability of a road network to resist, absorb and recover from serious disruptions. Figure 1.1 intuitively compares the three concepts of reliability, vulnerability and resilience and illustrates different levels of these concepts. $F(m)$, in Figure 1.1, refers to the ratio of the network performance at time m after perturbation to the equilibrium network performance. It is important to understand the impact of these events on network performance and how fast a network can recover to its normal operation to evaluate the vulnerability/resiliency of a road network. Moreover, in the case of evaluating network reliability, transportation planners and operators are often interested in understanding the connectivity, capacity, and service level of a network under long delays caused by recurrent and non-recurrent events because such occurrences can cause gridlock and, therefore, serious deterioration in network operation (Gu et al., 2019).

While the consequences of recurrent congestion have been studied in the literature, the important implications of non-recurrent occurrences on the performance of road networks and the nature of these extreme events demands more attention. In addition, the link between non-recurrent occurrences and travel time distribution, the impact on travel time variability, and their application in vulnerability analysis requires further investigation.

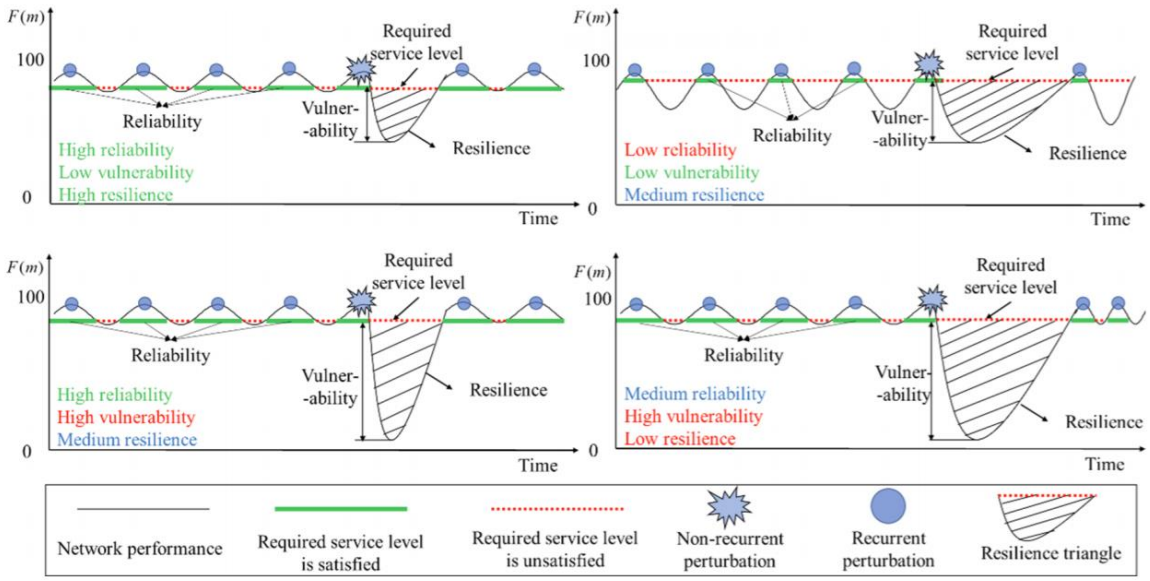

Figure 1.1**.** Illustration of the relationship between reliability, vulnerability and resilience (Gu et al., 2019)

While the probability of occurrences of non-recurrent occurrences is very low, the impact of these events on the performance of the transportation network can be huge. Therefore, in the transportation network, it is important to identify the vulnerable elements affected by such disruptive events. Such analysis helps to manage the potential risks resulting from these events and to better alleviate the disruptions to improve various aspects of transportation network performance. Vulnerable network elements (e.g., links, nodes) are part of a network responsible for a sharp decrease in traffic operation performance caused by capacity restrictions due to an incident or special event (Yperman & Tampere, 2006). Searching for vulnerable elements of a road network aims to identify potential critical elements in a transportation network. Once these elements are identified, network robustness can be improved by either reinforcing these elements or by improving alternative parallel routes (Matisziw and Murray, 2009).

Numerous evaluation approaches and techniques have been considered in the literature to identify the critical road infrastructure. The evaluation approach is strongly linked to how a transportation network is represented for the aim of the analysis. Depending on whether a transportation network is represented as an abstract graph or not, different measures of vulnerability are used in the literature. Apart from the evaluation approach, depending on how the scale of a disruptive event and its associated impact area is measured in a transportation network, different vulnerability analysis techniques were employed in the literature (Chen et al., 2012; Du et al, 2014; Jenelius and Mattsson, 2015; Taylor and Susilawati, 2012).

1.2 Motivating a more holistic approach to vulnerability analysis

While previously discussed techniques and approaches have been employed to measure the vulnerability of individual network components (e.g., links, nodes), in the presence of rare extreme events, these techniques identify the critical links of a transportation network measuring only the consequences of a link closure with little consideration given to the probability of link closure. The probability of link closure or failure is important as some of the links in a transportation network are more susceptible to rare events than others. Instances of such links are links with problematic geometric design, major links with higher fluctuations in travel demand, and/or major highways with high merging, diverging, and weaving traffic activity, which makes highways more prone to bottleneck formations and, thus, traffic breakdowns and possibly collisions. Moreover, depending on their type, rare events have a different probability of occurrences and their impacts differ case by case.

The need of a more comprehensive approach was underlined in several previous studies. Berdiaca (2002), Chen et. al (2007), Erath et al. (2009) and Watling and Balijepalli (2012)

specified that vulnerability analysis should be regarded as a risk analysis. Defining transportation system vulnerability as society's risk of transportation system disruptions and degradations, Jenelius (2009), Jenelius and Mattsson (2015) and Mattsson and Jenelius (2015) further specified that a vulnerability scenario should be formalised as a ''triplet'': 1) a description, 2) the probability and 3) the consequences (measure of damage) of that scenario. Therefore, a more holistic vulnerability analysis that takes into consideration this "triplet" approach is needed to simultaneously incorporate the probability and the consequences of disruptive events.

In addition, analytical vulnerability approaches discussed in the literature usually overlook important factors that affect the vulnerability of road links, such as extreme weather conditions, link specific geometry design, poor downstream traffic signal timing and/or links that are prone to a higher frequency of collisions. Multi-year observed travel time data carry important information that also reflect the historical spatial and temporal occurrence of link failure and their impact. This copious amount of multi-year travel time data can be further explored to identify the occurrence of disruptive events, examine the probability of occurrence of resulting extreme travel delays and evaluate their impact on overall network performance.

1.3 Objective and scope

This thesis studies road network vulnerability with consideration of both probability and consequences of rare events. Two vulnerability analysis sub-problems are examined: 1) road network vulnerability analysis by modeling monthly and seasonal variations in extreme travel delays, and 2) road network vulnerability analysis by measuring the spatiotemporal impact of incidents. The following is a brief overview of each sub-problem.

1.3.1 Introduction to the vulnerability analysis by modeling monthly and seasonal extreme travel delay variations

To the best of my knowledge, the impact of monthly and seasonal changes in weather condition on the variation of the extreme travel delay at the network level received less attention in network vulnerability research. The lack of studies on extreme travel time and the fact that the classic central theorem and traditional distributions fail to model extreme events motivates the development of a powerful modeling tool capable of capturing the variation in extreme travel delay data. My work attempts to fill the void in the literature by introducing a new class of extreme value distribution and proposes a new data-driven vulnerability approach, which considers both monthly and seasonal variation in extreme travel delay per km (also known as extreme normalized travel delay or extreme travel pace defined as the ratio of link travel time to link distance).

The risk factor of collisions as a function of drivers' perceptions leading to injuries or fatalities was estimated as 9 on snowy roads and 24 on icy roads compared to bare roads (Kilpeläinen and Summala, 2007). In addition, both the mean and SD of the observed speed decreases on slippery/icy road conditions (Rama and Kulmala, 2000; Saastamoinen, 1993). Because of the change in driving behavior, which itself varies according to road/weather conditions, the average observed extreme travel time, i.e., mean extreme travel time, varies seasonally and the variation in the observed extreme travel time, i.e., SD of extreme travel time, differs from season to season.

Of particular importance to the City of Calgary road network, which is the case study examined in this thesis, monthly and seasonal variation of extreme travel time is of interest as the frequency and severity of extreme events are highly correlated to the variation in weather conditions as an extrinsic cause of incidents (Alberta Transportation, 2016). People living in

southern Alberta face on average 55 snowy days every year (Environment Canada, 2019). The snow may remain on roads for even longer, especially during the fall and winter seasons when the temperature is usually below 0 degrees Celsius. Seasonal changes in weather contribute to extreme travel time variability. Snow and ice are contributors to 14% of fatal and 16.1% percent of non-fatal injury collisions (Alberta Transportation, 2016). In Finland, which has similar winter weather conditions, it has been reported that the probability of injury related collisions is over 20 times higher on an icy/snowy road than on a dry road surface (Kilpeläinen and Summala, 2007; Rama and Kulmala, 2000), resulting in long travel delays on affected links.

1.3.2 Introduction to the vulnerability analysis by measuring the spatiotemporal impact of incidents

Previous vulnerability analysis measures the increase in the total network travel cost associated with each individual link closure to accordingly determine the global vulnerability ranking. In the analysis conducted in the literature, the impact of each link closure is measured before and after closure by running a network traffic assignment without taking into account the duration of the closure and how quickly a link recovers from the incident or closure. In other words, it is not realistic to account for the impact of link closure by solely measuring the maximum possible increase in the travel time. This single point in time measure, in terms of increase in the travel time only when the link is fully closed, does not fully consider the dynamic ramifications of closure during the full-time span from link closure to recovery. This analysis is of interest as some of the links are more prone to incidents and simultaneously recover much more slowly from the incident compared to other links due to geometrical properties of the network in the vicinity of these links. Instance of geometrical properties influencing such analysis include lower network

redundancy in the vicinity of the links subject to incident, higher link betweenness, lack of alternative links to the links subject to incident, one directional links subject to incident, etc. Multi-year observed travel time and incident (i.e., collision and link closure) data carry this information and can be mined to determine the full spatial and temporal impact of link disruptions in the studied network.

Archived information about incidents, such as reported incident start time and clearance time, is usually based on the perceptions of those involved in the incident or guesswork by the officer or operator, thus, reported incident information is not accurate enough to rely on to determine the vulnerability of links to the presence of incidents. The true spatiotemporal impact of incidents could be used to determine the probability that the neighboring links of a link subject to incident (i.e., a link which is impacted directly by an incident) are affected by the incident and also the impact of the incident in terms of the magnitude of the increase in the pace of the neighboring links. Subsequently, network vulnerability ranking could take into consideration both the probability of neighboring links being affected by the incident as well as the magnitude of the impact.

In this thesis, I propose a new spatiotemporal incident impact determination approach by focusing on the similarity of congestion patterns in the vicinity of target links during an incident. In addition, I develop a risk-based vulnerability index which takes into consideration the probability of impact as well as the magnitude of the impact resulting from incidents.

1.4 Proposed methodology and research contributions

This research contributes to the body of knowledge by incorporating the probability and consequences of link failure in the road network vulnerability problem. Consideration for the

probability of the impact resulting from link failure is often overlooked in the literature even though this probability can significantly affect the order of the links in the network vulnerability ranking. The large size and the frequently high congestion levels of actual transportation networks further requires consideration of this probability.

The contribution of this study can be summarized in two distinctive categories. The contributions of this research to each category are outlined below.

Contributions to the network vulnerability while accounting for seasonality: The model developed for this research offers several contributions. The contributions are listed as follows:

- Developing a new class of extreme value distribution called compound generalized extreme value (CGEV) distribution which accounts for monthly and seasonal variation in the extreme travel delay.
- Proposing a new data-driven vulnerability approach that accounts for both the stochastic nature of extreme events, which is connected to the probability of extreme events, and link importance, which is used as an alternative measure for the relative impact of a link failure on network performance.
- Evaluating the performance of the proposed CGEV distribution and demonstrating that the CGEV distribution outperforms the traditional generalized extreme value (GEV) distribution in modeling the variation in the extreme travel delay.

Contributions to the network vulnerability while accounting for spatiotemporal impact of incidents: In this work, I attempt to address the existing gap in the literature through three main contributions. The contributions are outlined as follows:

- Developing a new methodology to determine the true spatial and temporal impact of incidents. I achieved this by monitoring the traffic pattern in the vicinity of the target link, also called impact area, and looking for similarities in the congestion pattern from which the temporal and spatial dimensions of the incident impact is determined.
- Developing a novel data-driven vulnerability index which takes into account the probability that links in the impact area associated with a given incident are affected. In addition, the consequence of the incident is measured in term of the deviation between the mean of the historical distribution of pace and the experienced pace during the unusual event of that link.
- Modeling the vulnerability of large size networks by applying real observed data from the Calgary road network.

1.5 Thesis organization

This thesis consists of four chapters that are laid out as follows:

Chapter 2 is devoted to the derivation of the CGEV distribution by accounting for monthly and seasonal variations in the extreme travel delay resulting from incidents. This chapter begins with a comprehensive overview of previous studies on the relationship between extreme events and travel time distribution, travel time variability and road network vulnerability, followed by derivation of the CGEV distribution. This chapter also covers problem formulation for the vulnerability index and a real network case study (Calgary road network).

Chapter 3 describes a data-driven vulnerability approach accounting for spatiotemporal impact of historical incidents. This chapter includes the formulation of the proposed model and several examples that demonstrate its applicability. The chapter also includes formulation of the proposed data-driven vulnerability index and the vulnerability analysis of Calgary road network.

Chapter 4 summarizes the findings of this research and concludes the work described in this dissertation. The contributions of this research to the greater body of literature are described and recommendation for future research are made.

CHAPTER 2: COMPOUND GENERALIZED EXTREME VALUE DISTRIBUTION FOR MODELING THE EFFECT OF MONTHLY AND SEASONAL VARIATION ON THE EXTREME TRAVEL DELAYS FOR VULNERABILITY ANALYSIS OF ROAD NETWORKS [1]

2.1 Background and related studies

2.1.1 Travel time distribution and extreme events

While traffic information is almost always communicated by means of average travel time, travelers usually experience variation in travel times from one time period to another during a day, from one day to another, and over different months and seasons. Thus, a stochastic representation of travel time and its variation is of more interest than specifying a deterministic value or range for the average travel time. Characterizing travel time variability is the basis of many travel time reliability analyses that resulted in numerous derived measures to represent the reliability of travel time (Asakura and Kashiwadani, 1991; Florida Department of Transpiration, 2000; Chen and Recker, 2000; Chen et al., 2003). State-of-the-art practices often use the mean and variance of travel time to derive buffer time and planning time indices as measures of road reliability. However, empirical travel time analyses have shown that travel time distributions are not necessarily symmetrical but highly skewed to the right with a heavy tail, especially in the presence of road network disruptions resulting from adverse weather conditions, car collisions, or other incidents (Bogers et al., 2006; Fosgerau and Karlström, 2010; Sumalee et al., 2013; Susilawati et

[1] The contents of this section have been used in the paper entitled: "Compound Generalized Extreme Value distribution for modeling the effects of monthly and seasonal variation on the extreme travel delays for vulnerability analysis of road network". Revise and resubmit in "Transportation Research Part C: Emerging Technologies"

al., 2013; and Van Lint et al., 2008). Common indices of buffer and planning time are thus not illuminating enough for evaluating heavy-tailed travel time distributions.

Different parametric distributions were tested in previous studies to describe the observed travel time data and its stochastic nature. Examples of such distributions include log-normal (Herman and Lam, 1974; Richardson and Taylor, 1978; Pu, 2011), truncated log-normal (Wang et al., 2012), gamma (Herman and Lam, 1974; Polus, 1979; Kim and Mahmassani, 2015), beta (Polus, 1979; Castillo et al., 2012), Weibull (Al-Deek and Emam, 2006), exponential (Talley and Becker, 1987; Noland and Small, 1995), bimodal (Yang et al., 2014), and Burr type XII (Susilawati et al., 2013). In a more recent study, Kim and Mahmassani (2015) presented a compound gamma-gamma distribution to describe the travel delay per mile assuming a multiplicative error model for both vehicle-to-vehicle and day-to-day travel time variability.

While travel time variability and travel time distribution have received increased attention in the literature, much less attention is paid to examine the tail of the travel time distribution, which consists of extreme travel times associated with the occurrence of recurrent and non-recurrent extreme events (e.g., incidents and extreme weather conditions). The impact of the extreme travel times due to both recurrent and non-recurrent events is of interest as the demand and supply interaction patterns resulting from these events can lead to long delays and, therefore, link failure. Extreme travel times are low-probability events located on the right-hand side tail of a distribution and can result in link failure and possibly gridlocks, which contribute to road network vulnerability. Fosgerau and Fukuda (2012) showed that traditional distributions, which are usually used to model travel time data, did not provide a good description of extreme travel time data. Xu et al. (2014) introduced a new risk-based measure of travel time reliability that characterized the tail of the travel time distribution. Zhong et al. (2019) introduced a double time-scale travel time

distribution not only to model travel time under risk-free conditions but also to capture the skewed distributions of travel time under abnormal traffic conditions. The authors further applied the proposed double time-scale travel time distribution to analyzing network resilience by showing the short-term and long-term impacts of disruption and systematic changes on road network performance.

2.1.2 Travel time variability and the treatment of extreme events

Different measures of travel time variability are considered in the literature such as vehicle-to-vehicle and period-to-period. Vehicle-to-vehicle variability refers to the variation in travel time of vehicles that depart within the same period of time. For example, morning peak and period-to-period variation in travel time refers to how the mean travel time within a specific time interval, for example, morning peak, varies over different periods of time (e.g., day-to-day, month-to-month, etc.) (Noland and Polak, 2002; Yildirimoglu et al., 2015). Numerous studies have investigated the relationship between the mean value and SD of travel time across vehicles and from one time period to another. The relationship between the mean and variance of travel time was first explored in the kinetic theory of traffic flow developed by Prigogine and Herman (1971). In another study, Jones et al. (1989) indicated that the mean and SD of normalized travel time were highly positively correlated. In studying the mean-SD travel time relationship, normalized travel time was first used by Herman and Lam (1974) to differentiate between travel time variability from heterogeneity in speed and variability due to trip distance. While vehicle-to-vehicle variability is the result of interactions between the vehicles, the period-to-period variability is mainly caused by extrinsic occurrences such as a change in weather conditions, incident occurrences, work zones, etc. (Cohen and Southworth, 1999; Kwon et al., 2000; Kwon et al.,

2011), and variation of capacity (Brilon et al., 2005); which itself is stochastic and can vary according to road surface conditions (wet or dry), outdoor light level (daylight or darkness), and the prevailing purpose of the road (long distance or metropolitan commuter traffic) (Ponzlet, 1996).

A more common period-to-period travel time variability addressed in the literature is day-to-day travel time variability. Different definitions are considered to describe the day-to-day variation in travel time; thus, various functional forms are suggested to fit the day-to-day travel time data. In an early study, Turner and Wardrop (1951) noted that the day-to-day mean-SD plot was well represented by a power function where the exponent of mean travel time was obtained as 0.68. Herman and Lam (1974) theoretically proved that the SD of the travel time was proportional to the square root of the mean travel time, assuming that the travel times on each route were uncorrelated and identically distributed. However, working with the empirical data, they indicated that the linear function provided a better fit, and the root-square function may not be valid for long journeys (travel time greater than 30 minutes). Using empirical bus travel time data, Polus (1979) obtained the square-root function for a mean-SD relationship. Based on the theoretical work of Herman and Lam (1974), the square-root relationship between the mean and SD of travel time was considered that from which the mean-coefficient of variation (CV) square root function was derived (Richardson and Taylor, 1978), and the square-root relationship represented the linear relationship between congestion factor- defined as mean travel time divided by free-flow speed and CV (Taylor, 1982). Using travel time data from London, UK, May et al. (1989) showed that both linear and square-root functions perform equally well.

While link travel time is stochastic and purely dynamic, mean travel time can be used as an explanatory variable to predict travel time variability. A few studies have made important

contributions in showing the existence of a linear relationship between mean travel time and SD of travel time (Mohammadi, 1997; Mazloumi et al., 2009; Hellinga, 2011; Kim and Mahmassani, 2015 and Kouwenhoven and Warffemius, 2016). In these studies, extreme travel time data are usually treated as outliers and accordingly disregarded from the data set as shown in Figure 2.1. Hellinga (2011) removed the extreme travel time data to obtain a linear relationship between the annual average travel time and SD. Similarly, de Jong and Bliemer (2015) and Kouwenhoven and Warffemius (2016) indicated that the extreme travel time data may affect the SD of travel time; hence, the extreme travel time data were filtered to obtain a linear function for the mean-SD plot. The occurrence of these extreme travel times is highly stochastic and hardly observed as they are often associated with rare extrinsic occurrences (e.g., special events, collisions, signal failures, extreme weather conditions, etc.), making it even more challenging to establish a meaningful relationship between the mean and SD of extreme travel times. Establishing a meaningful relationship between mean and SD of extreme travel times will consider the effect of long delays on network performance and improve the understanding of the statistical properties of extreme travel times. Better apprehension of the stochastic nature of extreme travel times will in turn aid in establishing proper vulnerability measures that consider the effect of long delays. Thus, extreme travel times will not be filtered from the analysis to gain insight into their statistical properties. Therefore, there is a need for a systematic way of quantifying, analyzing, and assessing the variability of extreme travel times, and then various vulnerability measures can be developed. This research takes many steps in this direction.

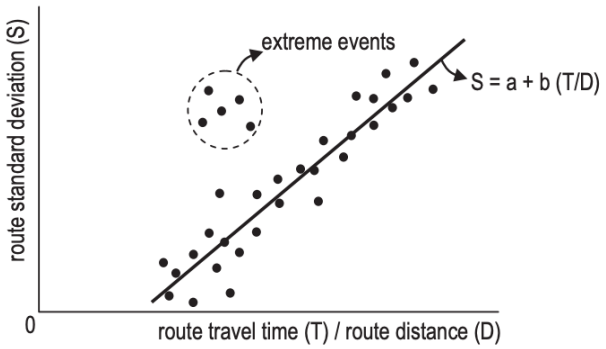

Figure 2.1. Relationship between mean and standard deviation of travel time in the presence of extreme events (Jong and Bliemer 2015)

To my knowledge, none of the published studies appears to account for the effect of variability in extreme travel time. The last two decades witnessed the emergence of floating probe data with large spatial network and temporal coverage. This multi-year network-wide observed extreme travel time data contain important information that reflect the spatial and temporal occurrence of recurrent and non-recurrent extreme travel delays. This data can further advance the understanding of the distribution and variability of extreme travel times and their effects.

2.1.3 Road network vulnerability and extreme events

Road network vulnerability analysis has recently received much attention in the literature. Depending on how a road network is represented in the vulnerability analysis problem, two distinct evaluation approaches exist with little overlap: 1) topological vulnerability analysis of road networks and 2) system-based vulnerability analysis of road networks. In the topological vulnerability analysis approach, a road network is typically represented as an abstract network (graph) consisting of nodes and undirected, unweighted links. Topological vulnerability studies usually deploy accessibility and network efficiency indices as measures of vulnerability to assess

the socioeconomic consequences of link closure. Examples of recent topological network vulnerability studies include important analysis conducted on road networks (Bell et al., 2017; Gao et al., 2019; López et al., 2017) and on public transport networks (Li et al., 2019; Ye and Kim, 2019). In contrast, system-based vulnerability studies employ graph representation of real networks consisting of weighted links with weights corresponding to actual link length, travel cost, etc. Transport network performance is examined through the interaction between travel demand and supply in terms of the increase in the total travel cost of a network to identify the critical elements of road networks (Almotahari and Yazici, 2019; García-Palomares et al., 2018) and public transport networks (Jiang et al., 2018; Lu, 2018; Yap et al., 2018).

As reported in previous studies, a more comprehensive approach to vulnerability analysis is required, and this new approach needs to be taken from a risk analysis perspective. However, due to the difficulty in predicting the probability of extreme event occurrences given the lack of historical network-wide travel information, the majority of the previous studies still approach the vulnerability problem by only focusing on the theoretical, consequential aspects of extreme events (Faturechi and Miller-Hooks, 2014; Taylor, 2017; Gu et al., 2019). Clearly, a comprehensive vulnerability analysis is needed to incorporate both the probability and the consequences of rare events.

2.2 Extreme value theory (EVT)

In the analysis in which the probability of the tail of the probability density function (PDF) is of interest, a classical central limit theorem may not be applicable. Consider a section of a freeway with n samples of independent and identically distributed (i.i.d.) random travel times

represented as X_1, X_2, …, X_n. The probability of a travel time exceeding a certain threshold x, can be then calculated by the following:

$$P(X \geq x) = 1 - F(x) \tag{2.1}$$

where F(x) is the cumulative distribution function (CDF) that follows a certain statistical distribution. This method may be used to estimate the probability of the tail of the travel time distribution for a relatively large value of x. However, in the case that the expected value of X, i.e., E(X), is not finite, the central limit theorem does not apply. Moreover, if the probability of the travel time exceeding the maximum observed travel time in the sample yields a value of 0, as represented mathematically below, the central limit theorem is no longer relevant:

$$p = P(X > x) = 1 - F(x) = 0 \qquad for\ x \geq x_{max} \tag{2.2}$$

where x_{max} is the maximum observed travel time. While the probability of these extreme travel times is very low, it cannot be assumed that such travel times are impossible, suggesting that special statistical techniques that focus on extreme events, or identically small tail probabilities, need to be developed.

The extreme value theory was first proposed by Fisher and Leonard Tippett (1928) and was further developed by Gnedenko (1943). Based on the Fisher–Tippett–Gnedenko theorem, if X_1, X_2, …, X_n are samples of i.i.d. random variables (travel time in this study), and M_n=max {X_1, X_2, …, X_n} denotes the maximum value among n sampled data, then the distribution of M_n converges to the following distribution called the GEV distribution:

$$H(x) = \begin{cases} exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & if\ \xi \neq 0 \\ exp\left[-exp\left(-\frac{x-\mu}{\sigma}\right)\right] & if\ \xi = 0 \end{cases} \quad (2.3)$$

where μ and σ denote location and scale parameters, respectively. ξ is called a shape parameter, and depending on the value ξ takes, $H(x)$ converges to different distributions. If $\xi \to 0$, H(x) corresponds to the Gumbel distribution (also called GEV distribution Type-I). If $\xi > 0$, H(x) corresponds to the Fréchet distribution (also called GEV distribution Type-II), and if $\xi < 0$, $H(x)$ corresponds to the Weibull distribution (also called GEV distribution Type-III).

Block maxima is the method of choosing the maximum travel time in each sample travel time and it is used in this study to determine the extreme travel times in an examined network. In this approach, the main data set is divided into equal blocks and the maximum value of each block is determined as a single member of the subset.

EVT can be applied in various fields of study. It is a powerful tool to model extreme events in environmental, structural, and financial research on topics such as air pollution (Smith, 1989; Kütchenhoff and Thamerus, 1996; Ercelebi and Toros, 2009), climate change (Katz, 1999; Beniston et al., 2007; Cooley, 2009), wind severity classification (Lombardo et al., 2009), wind data analysis (Harris, 1996; Mayne, 1979; Holmes and Moriarty, 1999), windstorm losses (Rootzén and Tajvidi, 1997), temperature analysis (Ferrez et al., 2011), forest fires (Alvarado et al., 1998), geological and seismic studies [earthquake severity analysis (Campbell, 1982; Esfeh et al., 2016), earthquake risk analysis (Cornell, 1968), diamond data analysis (Caers et al., 1996)], hydrology [flood (Rossi et al., 1984; Frances et al., 1994; Renard and Lang, 2007; Willems et al., 2007), rainfall (Beguería and Vicente-Serrano, 2006; Papalexiou and Koutsoyiannis, 2013), drought (Bordi et al., 2007; Burke et al., 2010), wave modeling (Moeini et al., 2010)], financial

applications and risk management (McNeil, 1997; McNeil et al., 2000; Gencay et al., 2004), structural safety (Makkonen, 2008), corrosion (Rivas et al., 2008), and athletic records (Gembris et al., 2007).

However, the application of EVT is rather limited in transportation-related studies. One of the main transportation related applications of EVT is to model the unobserved portion of the utility function (error term) in the logit discrete choice behavior model (Luce and Suppes, 1965) and in a family of generalized extreme value models such as nested logit (Williams, 1977) and GEV (McFadden, 1978). EVT is also applied widely in road safety studies (Tarko and Songchitruksa, 2005; Zheng et al., 2014). EVT was also employed to estimate road traffic capacity (Hyde and Wright, 1986; Minderhoud et al., 1997) and to predict the violation of air quality standards at urban intersections (Sharma et al., 1999).

2.3 Proportionality between the mean and SD of extreme delays

Previous studies confirmed a highly positive correlation between SD of travel time and mean travel time for both within-day and day-to-day levels (Kim and Mahmassani, 2015). However, the relationship between the SD and mean value of extreme travel time has not yet been investigated. Observed travel time data over a period of 6 years was obtained from INRIX for the City of Calgary and was analyzed to investigate whether the SD and mean value of extreme travel data were correlated or not. INRIX collects roadway speed and travel time data from real-time anonymous phones, vehicles and mobile devices equipped with GPS locator devices and traditional road sensors and hundreds of other sources and aggregates them over 1-minute time intervals. INRIX also provides the segment length and free-flow travel time. The travel times of road segments were calculated given the roadway speed and length.

To reveal the relationship between the mean and SD of extreme travel time, travel time data from 6 years between the end of December 2013 to the end of December 2019 were collected and aggregated over 1-minute time intervals from a 0.82 km section of Shaganappi Trail NW N, a two-lane highway with a speed limit of 70 km/h. The block maxima method was applied to choose the extreme travel times out of the collected travel time data. The sample size of the block maxima approach was determined by the block intervals τ. Block intervals should be selected carefully to ensure a sufficiently large sample size. Choosing a block interval of τ results in a $\frac{N}{\tau}$ sample size, where N is the total number of days. $\frac{N}{\tau}$ also represents the total number of blocks or the equivalent total number of extreme travel times during the studied time span. For example, choosing a 1-day block interval results in 2,191 extreme travel time observations for 6 years. If a block interval is shorter, the sample size is larger. Consequently, a larger sample size results in a least biased mean and SD of the extreme travel time as calculated for each time interval *t* with the length of *T* per hour ($T > \tau$). However, the block interval should be large enough to ensure that each block has a sufficient number of observations and allows the extreme travel time to be observed among a relatively large set of travel data. Choosing a relatively large block interval is justified by the requirement that the extracted extreme travel time from each block interval can be guaranteed to be a truly extreme value. The block interval of 7 days/1 week, is considered in this study as a compromise between the two constraints. Choosing 7-day block intervals provides a sample size of 313 extreme observations over a period of 6 years.

After constructing the extreme travel time database, an appropriate value must be assigned to the time interval *T* to evaluate the period-to-period variation (monthly and seasonal variation in this study) in travel time. The mean and SD of extreme travel time is calculated for each time interval. Choosing a time interval of *T* results in $\frac{N}{T}$ of mean-SD points for period-to-period and a

total of $\frac{N}{90}$ points for seasonal variation - assuming each season includes on average 90 days - for which the linearity of the fitted curve must be examined. Similar to the block interval, *T* should be assigned an appropriate value. A large value of time interval *T* enables us to calculate the mean and SD of a relatively large number of extreme travel observations, which reduces the bias in calculating the mean and SD in each time interval. However, *T* should be small enough to have a sufficient number of mean-SD points for a seasonal level as choosing a large time interval significantly reduces the mean-SD data points. *T* is considered as one month in this study to accommodate these two constraints. Therefore, $\left[\frac{2191}{30.42}\right] = 72$ and $\left[\frac{2191}{90}\right] = 24$ mean-SD points are obtained for monthly and seasonal levels, respectively, where 30.42 represents the average number of days in a month.

Assuming that a linear relationship holds at both monthly and seasonal levels, the linear relationship between the mean and SD of extreme travel time data at the monthly level is expressed as follows:

$$\sigma_\epsilon = \theta_1 \mu_\epsilon + \theta_2 \quad (2.4)$$

Where:

ϵ- extreme travel time per unit distance (min/km)

σ_ϵ- SD of ϵ (min/km)

μ_ϵ- mean of ϵ (min/km)

θ_1, θ_2- coefficients

Equation (2.4) suggests that by increasing the mean of the extreme travel time, the SD of the extreme travel time also increases. When there is no variation in travel time, all vehicles most

likely experience only slight variation in the observed mean extreme travel time. The x-intercept as obtained by setting the SD in Equation (2.4) to zero is called the minimum observed mean extreme travel time, i.e., 0 mean extreme travel delay[2] as indicated in Figure 2.1. Therefore, the mean extreme travel time becomes equal to $(-\theta_2/\theta_1)$. Therefore, the x-intercept in the monthly plot represents the minimum observed mean extreme travel time.

The extreme travel delay can be calculated by subtracting the minimum observed mean of extreme travel time from each extreme travel time. It can be mathematically shown by $d = \epsilon - (-\theta_2/\theta_1)$ where d is a random variable that represents extreme travel delay per km. The mean and SD of extreme travel delay are $\mu_d = \mu_\epsilon + (\theta_2/\theta_1)$ and $\sigma_d = \sigma_\epsilon$, respectively. μ_ϵ can be written as $\mu_\epsilon = \mu_d - (\theta_2/\theta_1)$. Equation (2.5) is obtained by replacing μ_ϵ and σ_ϵ in Equation (2.4) as follows:

$$\sigma_d = \theta_1 \mu_d \tag{2.5}$$

In Equation (2.5), θ_1 is the coefficient of variation (CV) of the extreme travel delay, and the x-intercept of the monthly extreme travel delay data is 0.

For time intervals *t=1, …, M* in each season *j=1, …, S*, the mean and SD of extreme travel delay data can be obtained as follows:

$$\mu_{d_{tj}} = \frac{1}{n_{tj}} \sum_{i=1}^{n_{tj}} d_{itj} \tag{2.6}$$

[2] This extreme travel delay is obtained by subtracting the x-intercept (i.e., the minimum observed mean extreme travel time) from the extreme travel time and is different from the actual extreme travel delay, which is obtained by subtracting the extreme travel time from the free-flow travel time. For the rest of the chapter, I refer to the former simply as *extreme travel delay*. When needed, the latter will be indicated as *actual travel delay*. The extreme travel delay can be converted into the actual travel delay anytime given the x-intercept and the free-flow travel time.

$$\sigma_{d_{tj}} = \sqrt{\frac{1}{n_{tj}} \sum_{i=1}^{n_{tj}} ({d_{itj}}^2 - {\mu_{d_{tj}}}^2)} \quad (2.7)$$

Where:

i- block interval index

t- time interval index

j- season index

M- total number of time intervals in each season

S- total number of seasons

$\mu_{d_{tj}}$- mean of extreme travel delay at time interval *t* in season *j*

$\sigma_{d_{tj}}$- SD of extreme travel delay at time interval *t* in season *j*

n_{tj}- total number of blocks at time interval *t* in season *j*

d_{itj}- extreme travel delay of block i at time interval t in season *j*

The linear relationship between the mean and SD of extreme travel delay can be rewritten:

$$\sigma_{d_{tj}} = \theta \mu_{d_{tj}} \quad (2.8)$$

where θ represents the CV of the d_{tj}.

Similar to the monthly variation in extreme travel delay, the mean and SD of the seasonal extreme travel delay variation are calculated assuming D_{tj} as the mean extreme travel delay for time interval *t* in season *j*.

$$\mu_{D_j} = \frac{1}{M}\sum_{t=1}^{M} D_{tj} \tag{2.9}$$

$$\sigma_{D_j} = \sqrt{\frac{1}{M}\sum_{t=1}^{M}({D_{tj}}^2 - {\mu_{D_j}}^2)} \tag{2.10}$$

Where:

μ_{D_j}- mean extreme travel delay in season *j*

σ_{D_j}- SD of extreme travel delay in season *j*

In Equation (2.11), D_{tj} corresponds to $\mu_{d_{tj}}$ in Equation (2.6) as follows:

$$D_{tj} = \mu_{d_{tj}} = \frac{1}{n_{tj}}\sum_{i=1}^{n_{tj}} d_{itj} \tag{2.11}$$

The linear relationship between the mean and SD of seasonal extreme travel delay variation can be expressed as follows:

$$\sigma_{D_j} = \Theta\mu_{D_j} \tag{2.12}$$

where Θ denotes the CV of D_j.

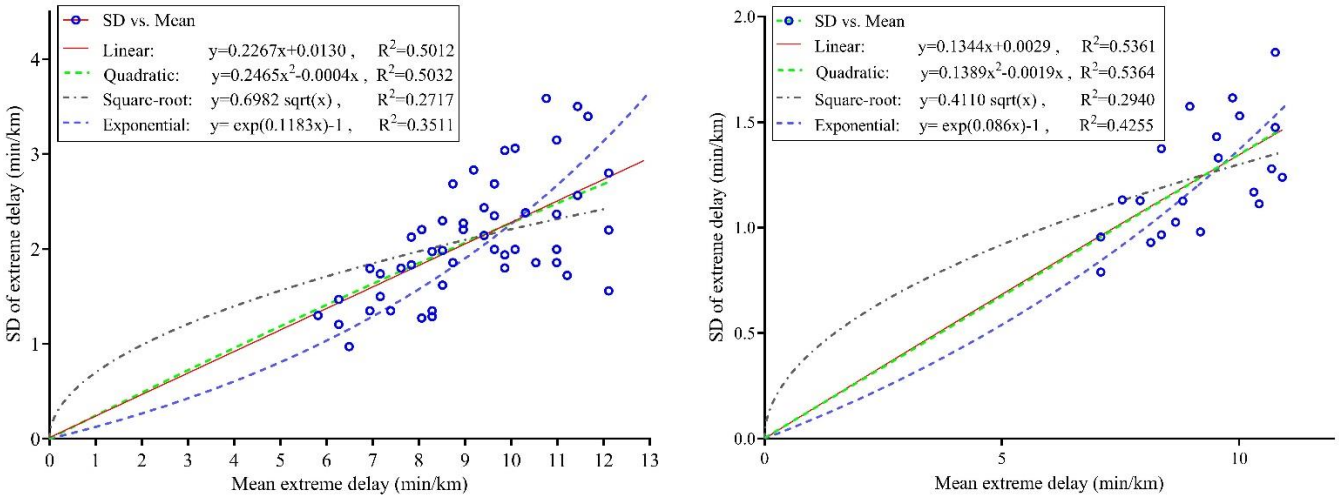

Figure 2.2. Monthly (left) and Seasonal (right) mean-SD plots for extreme travel delay on a road segment in Calgary

To validate the linear relationship between the mean and SD of extreme travel delay represented in Equations (2.8) and (2.12), a number of curves previously used in the literature to describe the mean-SD relationship of travel delay (linear, quadratic, exponential and square-root) are fitted to the monthly and seasonal extreme travel delay data for multiple links in the Calgary network. An example is depicted in Figure 2.2 for the 0.82 km section of Shaganappi Trail NW N. As can be seen in Figure 2.2, the SD of the extreme travel delay increases as the mean of extreme travel delay increases. This finding is intuitive as similar trend has been found in the literature for vehicle-to-vehicle and period-to-period travel time variability (Mahmassani et al., 2012, Kim and Mahmassani, 2015). According to Prigogine and Herman (1971), the mean travel time and its variance changes according to traffic conditions. In uncongested traffic condition, drivers are able to maintain fairly high and relatively similar speed which contributes to the findings that relatively short mean travel times with low standard deviation prevail. On the other hand, increased traffic congestion causes decrease in the average speed of the drivers. Consequently, drivers experience increased interactions with other vehicles in their vicinities. These interactions lead to increase in

the fluctuations of the traffic speed which consequently causes increase in the SD of travel time (Mahmassani et al., 2012).

While the R-square value obtained from fitting the quadratic curve is slightly better in both monthly and seasonal variations in extreme travel delay compared to the linear model (0.5032 vs. 0.5012 in the monthly case and 0.5364 vs. 0.5361 for the seasonal case), the quadratic model tends to overfit both monthly and seasonal data; a slight improvement in the associated R-square values is not statistically significant and cannot be justified because it is an overly complex model. Moreover, a simple linear model results in a less complex representation of the model, which is desirable for describing the period-to-period and seasonal variations of extreme travel delay. As shown in Figure 2.2, the intercepts of the obtained linear relationship between the mean and SD of extreme travel delay for both period-to-period and seasonal levels (-0.058 and -0.022, respectively) are not significantly different from 0 for a 99% confidence interval given the t-stat value of 0.035 and 0.008, respectively, which further verifies the linear relationships represented in Equations (2.8) and (2.12). These linear relationships are validated using data from other links in the Calgary network with different lengths and speed limits. The average R-square value is found as 0.71 with the SD of 0.17 calculated across 5,104 links in Calgary road network. In brief, the linear model is better in modeling the monthly variation in extreme travel delay compared to the quadratic, exponential and square root functions. Moreover, the linear model represents a good fit to the seasonal variation in extreme travel delay data. The results confirm that a linear relationship exists between the mean and SD of extreme travel delay data, which is valid for both the monthly and seasonal levels.

2.4 Extreme travel delay modeling

2.4.1 Multiplicative error model

In this section, the individual extreme travel delay on a given link is modeled. The extreme travel delays that belong to time interval *t* are denoted by d_t (min/km). As shown in Figure 2.2, some degree of heteroscedasticity exists in the monthly variation in the observed extreme travel time mean-SD data. The observed heteroscedasticity requires a modeling tool capable of addressing the time-varying variance resulting from the observed heteroscedasticity. In addition, using Equation (2.7) to calculate the SD of extreme travel delay for different time intervals implies that the variance of extreme travel delay data is not constant and varies over time, as shown in Figure 2.2, where the SD of the individual extreme travel delay data varies between 0.97 and 3.60 (min/km) on the y-axis. A varying SD suggests that the error term of the individual extreme travel delay cannot be expressed by an additive error model in which the variance of the error component is assumed to be constant. A multiplicative error model is utilized to address the variability of the extreme travel delay variance; the model describes the relationship between an individual extreme travel delay and its mean as shown in Equation (2.13):

$$d_t = D_t . \varepsilon_d \tag{2.13}$$

where D_t is the mean of d_t, which represents the mean extreme travel delay at time interval *t*, and ε_d is the error term, which denotes the period-to-period variation of the extreme travel delay around its mean.

Given positive values of d_t (extreme travel delay) and D_t (mean extreme travel delay), the error term in Equation (2.13) only takes positive values. In addition, assuming a multiplicative error model, the expected value of ε_d, denoted by $E(\varepsilon_d)$, must be equal to 1 (Engle, 2002). The

random error term ε_d can follow any continuous PDF that satisfies the above specifications. Examples of such distributions include the exponential, gamma, log-normal, log-logistic, Pareto, and GEV distributions. Choosing the best-fitted distribution to ε_d requires identifying which PDF best describes the period-to-period individual extreme travel delay (d_t). In this study, since the observed extreme travel delay data (d_t) follow the GEV distributions as discussed in detail in section 2.2, the error term (ε_d) is assumed to follow the GEV distribution with a unit mean.

More specifically, the error term follow the GEV distribution with location parameter μ_d, shape parameter ξ_d, and scale parameter $\frac{\xi_d(1-\mu_d)}{\Gamma(1-\xi_d)-1}$ [3] denoted by $\varepsilon_d \sim GEV\left(\mu_d, \frac{\xi_d(1-\mu_d)}{\Gamma(1-\xi_d)-1}, \xi_d\right)$, where $\Gamma(\cdot)$ is the gamma function. Consequently, d_t in Equation (2.13) follows the GEV distribution denoted by $d_t \sim GEV\left(\mu_d D_t, \frac{\xi_d(1-\mu_d)D_t}{\Gamma(1-\xi_d)-1}, \xi_d\right)$.

In the next step, the seasonal variation in extreme travel delay is modeled also using a multiplicative error model as follows:

$$D_t = \Delta_j . \varepsilon_D \tag{2.14}$$

where Δ_j is the mean of D_t and represents seasonal extreme travel delay considering time interval of length *T*. ε_D is the error component and represents the seasonal variation of the mean extreme travel delay around Δ_j. Similar to the period-to-period case, ε_D is assumed to follow the GEV distribution with a unit mean. Similar to d_t, the extreme value theory can also apply to D_t, which is the mean of d_t.

[3] Please refer to Appendix I on how the scale parameter is calculated based on the location parameter and shape parameter.

Proposition 1. Assume d_{1tj}, d_{2tj}, …, d_{itj} are a series of period-to-period extreme travel delay observations where d_{itj} is the extreme travel delay of block *i* at time interval t for season *j*. Also, assume that D_{tj} is the seasonal mean of the extreme travel delay of time interval *t* and season *j* where $D_{tj} = \frac{1}{n_{tj}} \sum_{i=1}^{n_{tj}} d_{itj}$. The random variable D_{tj}, which represents the period-to-period mean of extreme travel delay (D_{tj}), follows the GEV.

Proof 1. From the definition of the period-to-period extreme travel delay, Equation (2.15) can be used to describe d_{itj}.

$$d_{itj} = max\{d_{1itj}, d_{2itj}, \dots, d_{hitj}\} \tag{2.15}$$

where d_{hitj} denotes h[th] travel delays of block *i* at time interval t in season *j*. Equation (2.15) states that in each block, the maximum individual delay is chosen as the extreme travel delay. Choosing any other individual travel delay rather than the maximum in block 1 results in a different period-to-period mean of extreme travel delay, which can be shown by D_{tj}^1 where $D_{tj}^1 \leq D_{tj}$. Similarly, the period-to-period mean of extreme travel delay that is obtained by choosing an individual travel delay rather than the maximum in block *i* results in D_{tj}^i where $D_{tj}^i \leq D_{tj}$. Thus, in a series $\{D_{tj}^1, D_{tj}^2, \dots, D_{tj}^i, \dots, D_{tj}, \dots, D_{tj}^n\}$, D_{tj} is the maximum (extreme) period-to-period mean of extreme travel delay in time interval *t* in season *j*; thus, the extreme value theory is applied to D_{tj}.

Since D_t follows the GEV, the error term ε_D in Equation (2.14) also follows the GEV. More specifically ε_D follows the GEV with location parameter μ_D, shape parameter ξ_D, and scale parameter $\frac{\xi_D(1-\mu_D)}{\Gamma(1-\xi_D)-1}$ denoted by $\varepsilon_D \sim GEV\left(\mu_D, \frac{\xi_D(1-\mu_D)}{\Gamma(1-\xi_D)-1}, \xi_D\right)$. Since the extreme travel delay distribution is usually bounded on the left side and highly skewed to the right, it is likely that both

$\xi_d \neq 0$ and $\xi_D \neq 0$. In the case that any shape parameter is equal to 0, the associated GEV distribution will not have a lower bound, which is unlikely in the case of extreme travel delay distribution. Thus, it is assumed that $\xi_d \neq 0$ and $\xi_D \neq 0$, and therefore, the first part of the GEV distribution in Equation (2.3) applies to both period-to-period and seasonal extreme travel delay data. Table 2.1 summarizes the specifications of the discussed models. For the period-to-period distribution model, the CV of the error term is equal to the CV of the extreme travel delay, which is θ in Equation (2.8). Similarly, for the seasonal distribution model, the CV of the error term is equal to the CV of the mean extreme travel delay, which is Θ in Equation (2.12). The CVs of both the extreme travel delay and mean extreme travel delay can be used to compare the CV enhanced by the GEV distributions with the slopes of the linear models fitted to the period-to-period (θ) and seasonal extreme travel delay data (Θ). This comparison provides more insight into how well the individual GEV distribution fits the empirical data.

Table 2.1. Characterizing period-to-period and seasonal extreme travel delay variability

	Period-to-period distribution	Seasonal distribution
Error model	$d_t = D_t . \varepsilon_d$	$D_t = \Delta_j . \varepsilon_D$
Distribution	GEV	GEV
Properties of error term distribution	$\varepsilon_d \sim GEV\left(\mu_d, \frac{\xi_d(1-\mu_d)}{\Gamma(1-\xi_d)-1}, \xi_d\right)$ $E(\varepsilon_d) = \mu_d + \frac{\xi_d(1-\mu_d)}{\Gamma(1-\xi_d)-1}\frac{\Gamma(1-\xi_d)-1}{\xi_d} = 1$ $VAR(\varepsilon_d) = \left(\frac{1-\mu_d}{\Gamma(1-\xi_d)-1}\right)^2 (\Gamma(1-2\xi_d) - \Gamma(1-\xi_d)^2)$ $CV(\varepsilon_d) = \frac{1-\mu_d}{\Gamma(1-\xi_d)-1}\sqrt{(\Gamma(1-2\xi_d) - \Gamma(1-\xi_d)^2)}$	$\varepsilon_D \sim GEV\left(\mu_D, \frac{\xi_D(1-\mu_D)}{\Gamma(1-\xi_D)-1}, \xi_D\right)$ $E(\varepsilon_D) = \mu_D + \frac{\xi_D(1-\mu_D)}{\Gamma(1-\xi_D)-1}\frac{\Gamma(1-\xi_D)-1}{\xi_D} = 1$ $VAR(\varepsilon_D) = \left(\frac{1-\mu_D}{\Gamma(1-\xi_D)-1}\right)^2 (\Gamma(1-2\xi_D) - \Gamma(1-\xi_D)^2)$ $CV(\varepsilon_D) = \frac{1-\mu_D}{\Gamma(1-\xi_D)-1}\sqrt{(\Gamma(1-2\xi_D) - \Gamma(1-\xi_D)^2)}$
Properties of response variable distribution	$d_t \sim GEV\left(\mu_d D_t, \frac{\xi_d(1-\mu_d)D_t}{\Gamma(1-\xi_d)-1}, \xi_d\right)$ $E(d_t) = D_t$ $VAR(d_t) = {D_t}^2\left(\frac{1-\mu_d}{\Gamma(1-\xi_d)-1}\right)^2 (\Gamma(1-2\xi_d) - \Gamma(1-\xi_d)^2)$ $CV(d_t) = \frac{1-\mu_d}{\Gamma(1-\xi_d)-1}\sqrt{(\Gamma(1-2\xi_d) - \Gamma(1-\xi_d)^2)}$	$D_t \sim GEV\left(\mu_D \Delta_j, \frac{\xi_D(1-\mu_D)\Delta_j}{\Gamma(1-\xi_D)-1}, \xi_D\right)$ $E(D_t) = \Delta_j$ $VAR(D_t) = {\Delta_j}^2\left(\frac{1-\mu_D}{\Gamma(1-\xi_D)-1}\right)^2 (\Gamma(1-2\xi_D) - \Gamma(1-\xi_D)^2)$ $CV(D_t) = \frac{1-\mu_D}{\Gamma(1-\xi_D)-1}\sqrt{(\Gamma(1-2\xi_D) - \Gamma(1-\xi_D)^2)}$

2.4.2 CGEV distribution to model extreme travel delay

A new equation can be obtained by combining Equations (2.13) and (2.14); this new equation describes the extreme travel delay (d_t) based on the two error terms as follows:

$$d_t = \Delta_j . \varepsilon_D . \varepsilon_d \tag{2.16}$$

Since d_t is the product of the multiplication of two randomly GEV distributed error terms, d_t can be modeled using a CGEV distribution. In this study, a new representation of CGEV distribution is derived as follows:

Let D_t follow the GEV distribution, i.e., $D_t \sim GEV\left(\mu_D \Delta_j, \frac{\xi_D(1-\mu_D)\Delta_j}{\Gamma(1-\xi_D)-1}, \xi_D\right)$. The PDF of D_t is given by the following equation:

$$f(D_t) = \frac{\Gamma(1-\xi_D)-1}{\xi_D(1-\mu_D)\Delta_j} exp\left[-\left(1+\frac{(D_t-\mu_D\Delta_j)(\Gamma(1-\xi_D)-1)}{(1-\mu_D)\Delta_j}\right)^{-\frac{1}{\xi_D}}\right].\left(1+\frac{(D_t-\mu_D\Delta_j)(\Gamma(1-\xi_D)-1)}{(1-\mu_D)\Delta_j}\right)^{-1-\frac{1}{\xi_D}} \tag{2.17}$$

The conditional PDF of d_t given its mean (D_t) under $d_t \sim GEV\left(\mu_d D_t, \frac{\xi_d(1-\mu_d)D_t}{\Gamma(1-\xi_d)-1}, \xi_d\right)$ is as follows:

$$f(d_t|D_t) = \frac{\Gamma(1-\xi_d)-1}{\xi_d(1-\mu_d)D_t} exp\left[-\left(1+\frac{(d_t-\mu_d D_t)(\Gamma(1-\xi_d)-1)}{(1-\mu_d)D_t}\right)^{-\frac{1}{\xi_d}}\right].\left(1+\frac{(d_t-\mu_d D_t)(\Gamma(1-\xi_d)-1)}{(1-\mu_d)D_t}\right)^{-1-\frac{1}{\xi_d}} \tag{2.18}$$

The unconditional PDF of d_t is obtained by integrating the multiplication of the PDFs in Equations (2.17) and (2.18) over all possible values of D_t from 0 to infinity.

$$f(d_t) = \int_0^\infty f(d_t|D_t) f(D_t)\, dD_t \tag{2.19}$$

The intermediate form of the PDF is as follows:

$$f(d_t) = \frac{\gamma_d \gamma_D}{\Delta_j \xi_d \xi_D} \int_0^{\infty} {D_t}^{-1} exp\left[-\left(1 - \mu_d \gamma_d + \gamma_d \frac{d_t}{D_t}\right)^{-\frac{1}{\xi_d}} - \left(1 - \mu_D \gamma_D + \gamma_D \frac{D_t}{\Delta_j}\right)^{-\frac{1}{\xi_D}} \right] \left(1 - \mu_d \gamma_d + \gamma_d \frac{d_t}{D_t}\right)^{-1-\frac{1}{\xi_d}} \left(1 - \mu_D \gamma_D + \gamma_D \frac{D_t}{\Delta_j}\right)^{-1-\frac{1}{\xi_D}} dD_t \quad (2.20)$$

Where $\gamma_k = \frac{\Gamma(1-\xi_k)-1}{1-\mu_k}$. Since there is no direct closed-form expression for the integral in Equation (2.20), I solve the problem using the Taylor series representation of the exponential function shown in Equation (2.21):

$$e^x = \sum_{v=0}^{\infty} \frac{x^v}{v!} \quad (2.21)$$

Expanding the exponential components in Equation (2.20) using Equation (2.21) results in Equation (2.22) as follows:

$$f(d_t) = \frac{\gamma_d \gamma_D}{\Delta_j \xi_d \xi_D} \int_0^{\infty} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} {D_t}^{-1} \frac{(-1)^{m+n}}{(m-1)!\,(n-1)!} \left(1 - \mu_d \gamma_d + \gamma_d \frac{d_t}{D_t}\right)^{-1-\frac{m}{\xi_d}} \left(1 - \mu_D \gamma_D + \gamma_D \frac{D_t}{\Delta_j}\right)^{-1-\frac{n}{\xi_D}} dD_t \quad (2.22)$$

Equation (2.22) can be approximated by Equation (2.23) considering the Taylor polynomial form of Equation (2.22), i.e., taking r initial terms of the Taylor series, and rearranging the integral in Equation (2.22):

$$f(d_t) = \frac{\gamma_d \gamma_D}{\Delta_j \xi_d \xi_D} \sum_{m=1}^{r} \sum_{n=1}^{r} \int_0^{\infty} {D_t}^{-1} \frac{(-1)^{m+n}}{(m-1)!\,(n-1)!} \left(1 - \mu_d \gamma_d + \gamma_d \frac{d_t}{D_t}\right)^{-1-\frac{m}{\xi_d}} \left(1 - \mu_D \gamma_D + \gamma_D \frac{D_t}{\Delta_j}\right)^{-1-\frac{n}{\xi_D}} dD_t \quad (2.23)$$

A closed-form algebraic expression based on a regularized hypergeometric function[4] is found for the integral in Equation (2.23) which results in Equation (2.24) as the closed-form expression for the CGEV distribution as follows:

$$f(d_t) = \frac{\gamma_d \gamma_D}{\Delta_j \xi_d \xi_D} \sum_{m=1}^{r} \sum_{n=1}^{r} \frac{(-1)^{m+n}}{(m-1)!\,(n-1)!} \frac{\pi\,(\gamma_d d_t)^{-1-\frac{m}{\xi_d}}\,(1-\mu_D\gamma_D)^{-1-\frac{n}{\xi_D}}\,csc\left[\pi(\frac{m}{\xi_d}-\frac{n}{\xi_D})\right]}{\Gamma\left(1+\frac{m}{\xi_d}\right)\Gamma\left(1+\frac{n}{\xi_D}\right)} f \tag{2.24}$$

Where f is as follows:

$$\begin{aligned} f = {} & \Gamma\left(1+\frac{n}{\xi_D}\right)^2 \left[\frac{\gamma_D}{\Delta_j(1-\mu_D\gamma_D)}\right]^{-1-\frac{n}{\xi_D}} \left[\frac{1-\mu_d\gamma_d}{\gamma_d d_t}\right]^{\frac{n}{\xi_D}-\frac{m}{\xi_d}} {}_2\tilde{F}_1\left(1+\frac{n}{\xi_D}, 1+\frac{n}{\xi_D}; \frac{n}{\xi_D}-\frac{m}{\xi_d}\right. \\ & \left. +1; \frac{\Delta_j\,(1-\mu_d\gamma_d)(1-\mu_D\gamma_D)}{\gamma_d\gamma_D d_t}\right) \\ & - \Gamma\left(1+\frac{m}{\xi_d}\right)^2 \left[\frac{\gamma_D}{\Delta_j(1-\mu_D\gamma_D)}\right]^{-1-\frac{m}{\xi_d}} {}_2\tilde{F}_1\left(1+\frac{m}{\xi_d}, 1+\frac{m}{\xi_d}; \frac{m}{\xi_d}-\frac{n}{\xi_D}\right. \\ & \left. +1; \frac{\Delta_j(1-\mu_d\gamma_d)(1-\mu_D\gamma_D)}{\gamma_d\gamma_D d_t}\right) \end{aligned} \tag{2.25}$$

Different values of r yield different approximations of the intermediate form of the PDF in Equation (2.20). For the ideal case of $r = \infty$, Equation (2.24) is the exact representation of Equation (2.20). However, calculating infinity terms of the polynomial Taylor expansion of the exponential function can be computationally exhaustive. The appropriate value of r used in the Taylor expansion to approximate the function depends on the desired tolerance. Figure 2.3 represents different PDFs with different parameters estimated for $r = 2, 3, 4, 5$ along with the intermediate form of the PDF. Considering a tolerance of 0.001, choosing the first five terms of

[4] Refer to Gradshteyn and Ryzhik (2014) for more information on the hypergeometric function and its properties.

the Taylor expansion ($r = 5$) of both exponential functions in Equation (2.20) provides a very close approximation of the intermediate form of the PDF, as shown in Figure 2.3.

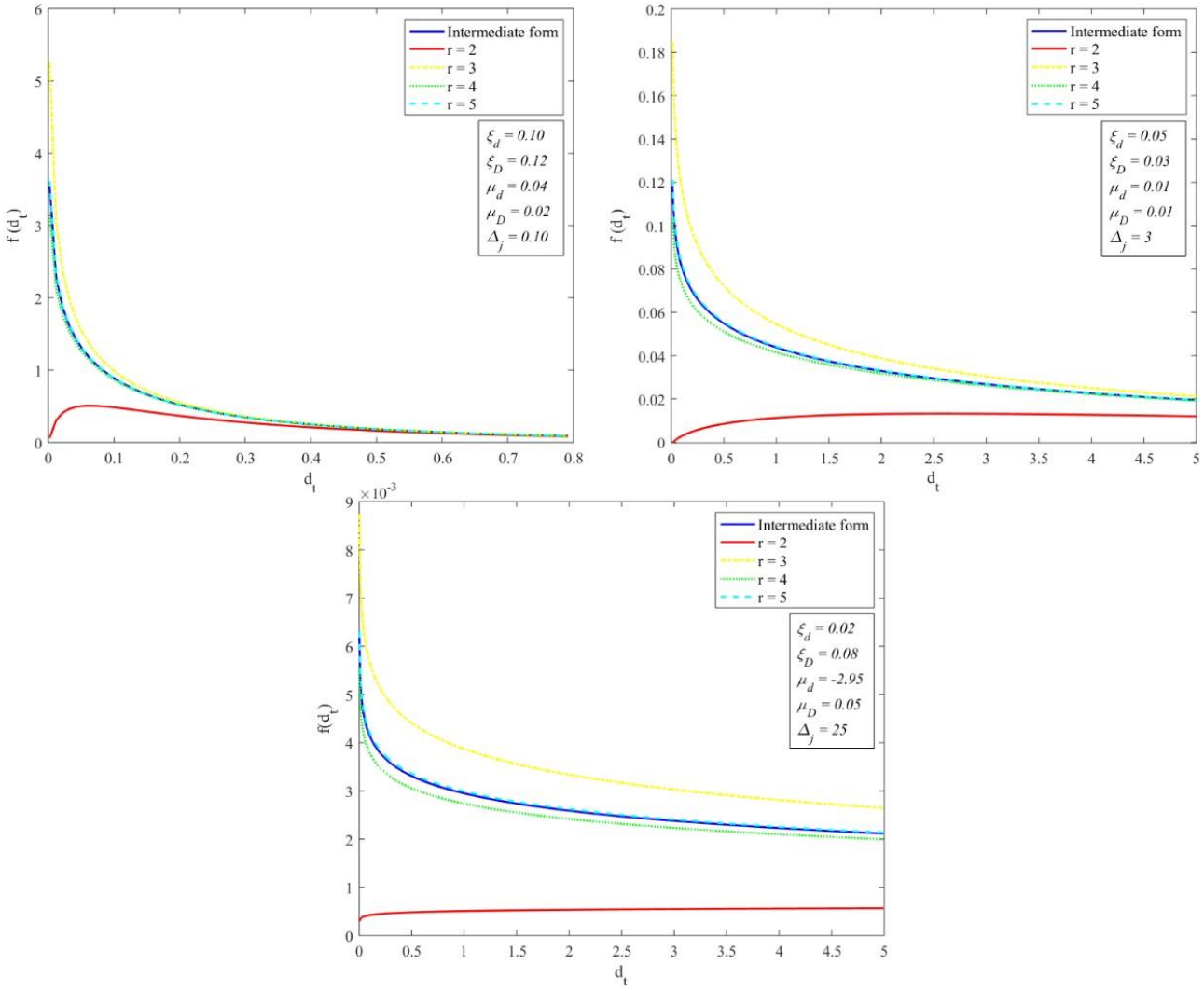

Figure 2.3. PDFs with different parameters approximated for different value of r versus the intermediate form ($\boldsymbol{d_t}$ is per $\boldsymbol{min/km}$)

The CDF of the CGEV distribution is then calculated by integrating the PDF as follows:

$$F(d_t) = \int_{-\infty}^{d_t} f(x)\ dx \tag{2.26}$$

The CDF of the CGEV distribution is obtained as follows:

$$F(d_t) = \frac{\gamma_d \gamma_D}{\Delta_j \xi_d \xi_D} \sum_{m=1}^{r} \sum_{n=1}^{r} \frac{(-1)^{m+n}}{(m-1)!\,(n-1)!} \frac{\pi\,(\gamma_d)^{-1-\frac{m}{\xi_d}} (1-\mu_D\gamma_D)^{-1-\frac{n}{\xi_D}} csc\left[\pi(\frac{m}{\xi_d}-\frac{n}{\xi_D})\right]}{\Gamma\left(1+\frac{m}{\xi_d}\right)\Gamma\left(1+\frac{n}{\xi_D}\right)} F \tag{2.27}$$

Where F is as follows:

$$\begin{aligned} F = {} & \Gamma\left(1+\frac{n}{\xi_D}\right)^2 \left[\frac{\gamma_D}{\Delta_j(1-\mu_D\gamma_D)}\right]^{-1-\frac{n}{\xi_D}} \left[\frac{1-\mu_d\gamma_d}{\gamma_d}\right]^{\frac{n}{\xi_D}-\frac{m}{\xi_d}} \frac{(d_t)^{-\frac{n}{\xi_D}}}{-\frac{n}{\xi_D}-n} \, {}_2\tilde{F}_1\left(1+\frac{n}{\xi_D}, 1+\frac{n}{\xi_D}; \frac{n}{\xi_D}-\frac{m}{\xi_d}\right. \\ & \left. +1; \frac{\Delta_j\,(1-\mu_d\gamma_d)(1-\mu_D\gamma_D)}{\gamma_d\gamma_D d_t}\right) \\ & -\Gamma\left(1+\frac{m}{\xi_d}\right)^2 \left[\frac{\gamma_D}{\Delta_j(1-\mu_D\gamma_D)}\right]^{-1-\frac{m}{\xi_d}} \frac{(d_t)^{-\frac{m}{\xi_D}}}{-\frac{m}{\xi_D}-m} \, {}_2\tilde{F}_1\left(1+\frac{m}{\xi_d}, 1+\frac{m}{\xi_d}; \frac{m}{\xi_d}-\frac{n}{\xi_D}\right. \\ & \left. +1; \frac{\Delta_j(1-\mu_d\gamma_d)(1-\mu_D\gamma_D)}{\gamma_d\gamma_D d_t}\right) \end{aligned} \tag{2.28}$$

The obtained CGEV distribution has five parameters, μ_d, ξ_d, μ_D, ξ_D, and Δ_j, where $\mu_d, \xi_d, \mu_D, \xi_D \in \mathbb{R}$ and $\Delta_j \in (0, \infty)$. Thus, the distribution of d_t is $d_t \sim \text{CGEV}(\mu_d, \mu_D; \xi_d, \xi_D; \Delta_j)$. μ_d and μ_D are the location parameters, ξ_d and ξ_D are the shape parameters, and Δ_j is the seasonal mean of the extreme travel delay in the CGEV distribution. μ_d and ξ_d represent the period-to-period variation in the extreme travel delay, and μ_D, ξ_D, and Δ_j characterize the seasonal distribution of extreme travel delay.

2.5 Analysis of extreme travel time data

2.5.1 Data

A real-world case study was conducted to showcase the procedures of calibrating the CGEV distribution and its applications in modeling the variability in extreme travel delay data

collected from various roadways in the City of Calgary over a long-term evaluation period. For this study, travel time data of the selected corridors were collected for each day between December 31, 2013 to December 31, 2019 for 6 years (2191 days); the data was aggregated over 1-minute time intervals to fully investigate the impact of rare events on the performance of the links. The update time interval of a minimum of 1 minute was required to have a sufficient number of travel times to determine the rare events. The data was collected by INRIX Roadway Analytics, which provides services that allow users to view and query regional traffic flow information. The coverage of the INRIX Roadway Analytics[5] includes provincial highways, the Major Road Network, and select arterial roads within the metro Calgary boundary.

The performance of the proposed approach is evaluated for the majority of the links in the Calgary road network. Around 85% of the Calgary road network, consisting of 5104 links, were selected for the analysis. To showcase the performance of the model at the link level, I selected a 0.82 km section of Shaganappi Trail NW N, Calgary as an example. This link connects several business centers and shopping centers with two residential areas. Moreover, this link is part of a corridor that connects two major roads in Calgary: 16th Avenue and Crowchild Trail. For the analysis, I divided the whole dataset into equal blocks of 1 week, i.e., τ = 7 days. Choosing block intervals of 1 week resulted in 313 extreme observations.

2.5.2 Investigating the linear relationship between the mean and SD of extreme travel delay

The linear relationship between the mean and SD of extreme travel delay is investigated by calculating the mean and SD of the extreme delays obtained from each block interval. The mean and SD of the extreme travel delay is calculated for both period-to-period (month-to-month) and

5 https://inrix.com/products/roadway-analytics/

seasonal levels using Eqs. (8) and (12), respectively. The linear model parameter estimation algorithm is summarized in terms of pseudocode as presented in Table 2.1.

Table 2.2. Algorithm for estimating the parameters of the mean-SD linear model

Procedure *LinearModelParametersEstimation*

Input: Historical travel time information for links in network G

Output: Parameters of the mean-SD linear model of extreme travel delay for each link

Initialization:

Assign an appropriate value to the block interval τ

Assign an appropriate value to the time interval T where $\tau \leq$ T

Step 1: Obtain x-intercept for period-to-period level

For each link $l \in G$

For each season $j=1, \ldots, S$

For each time interval $t=1, \ldots, M$

For each block interval $i = 1, \ldots, n_{tj}$

Set $t\_t_{itj} := max\{t\_t_{1itj}, t\_t_{2itj}, \ldots, t\_t_{hitj}\}$*

$\Longrightarrow \frac{\sum_{j=1}^{S}\sum_{t=1}^{M} n_{tj} = N}{\tau}$ blocks or the equivalent total number of extreme travel times

End for

Set $\mu_{t\_t_{tj}} := \frac{1}{n_{tj}}\sum_{i=1}^{n_{tj}} d_{itj}$, $\sigma_{t\_t_{tj}} := \sqrt{\frac{1}{n_{tj}}\sum_{i=1}^{n_{tj}}(t\_t_{itj}{}^2 - \mu_{t\_t_{tj}}{}^2)}$

End for

End for

$\Longrightarrow \frac{N}{T}$ mean-SD points for period-to-period variation in extreme travel times

Plot $\frac{N}{T}$ mean-SD points for period-to-period and obtain the x-intercept $(-\theta_2/\theta_1)$

Set $MMT_l := -\theta_2/\theta_1$**

Step 2: Obtain parameters of the linear models

For each season $j=1, \ldots, S$

For each time interval $t=1, \ldots, M$

For each block interval $i = 1, \ldots, n_{tj}$

Set $d_{hitj} := t\_t_{hitj} + \theta_2/\theta_1$

Set $d_{itj} := max\{d_{1itj}, d_{2itj}, \ldots, d_{hitj}\}$

$\Longrightarrow \frac{N}{\tau}$ total number of extreme travel delays

End for

Set $\mu_{d_{tj}} \coloneqq \frac{1}{n_{tj}} \sum_{i=1}^{n_{tj}} d_{itj}$, $\sigma_{d_{tj}} \coloneqq \sqrt{\frac{1}{n_{tj}} \sum_{i=1}^{n_{tj}} (d_{itj}{}^2 - \mu_{d_{tj}}{}^2)}$

End for

End for

$\Longrightarrow \frac{N}{T}$ mean-SD points for period-to-period variation in extreme travel times

Plot $\frac{N}{T}$ mean-SD points for period-to-period and obtain θ by performing linear regression

Set $\mu_{D_j} \coloneqq \frac{1}{M} \sum_{t=1}^{M} D_{tj}$, $\sigma_{D_j} \coloneqq \sqrt{\frac{1}{M} \sum_{t=1}^{M} (D_{tj}{}^2 - \mu_{D_j}{}^2)}$

$\Longrightarrow \frac{N}{T}$ mean-SD points for period-to-period variation in extreme travel delays

Plot $\frac{N}{90}$ mean-SD points for seasonal variation and obtain Θ by performing linear regression

End for

* t_t: travel time

‡ MMT: Minimum observed mean of extreme travel time

Time interval T should be selected such that enough period-to-period mean-SD extreme travel delay points exists for step 1 in Table 2.2 to estimate Θ. In this study, as indicated in Table 2.2, the minimum observed mean of extreme travel time is calibrated for each route segment separately. Calibrating the minimum observed mean of extreme travel time for each segment provides a more accurate estimation of extreme travel delay, especially when route segments are taken from the same corridor with different free-flow speeds.

2.5.3 Model calibration: estimating parameters of CGEV distributions

The CV of the error terms ε_d and ε_D obtained from fitting linear functions to both the extreme travel delay and mean extreme travel delay data, shown by θ and Θ and as calculated by

Eqs. (8) and (12), respectively, must be compared with the CVs enhanced by fitting CGEV distributions to both period-to-period and seasonal extreme travel delay data, shown by $\hat{\theta}$ and $\hat{\Theta}$, respectively, to validate the assumption that these error terms follow the GEV distribution. Based on the information provided in Table 2.1, the following equations are obtained to calculate $\hat{\theta}$ and $\hat{\Theta}$:

$$\hat{\theta} = \frac{1-\mu_d}{\Gamma(1-\xi_d)-1}\sqrt{(\Gamma(1-2\xi_d)-\Gamma(1-\xi_d)^2)} \tag{2.29}$$

$$\hat{\Theta} = \frac{1-\mu_D}{\Gamma(1-\xi_D)-1}\sqrt{(\Gamma(1-2\xi_D)-\Gamma(1-\xi_D)^2)} \tag{2.30}$$

For this purpose, the steps outlined in Table 2.3 are followed to validate the assumed distribution for the error terms at both link and network levels.

The parameters of the CGEV distribution are estimated by fitting 2 GEV distributions to the error terms ε_d and ε_D using the maximum likelihood estimation (MLE) method. This estimation results in 2 shape parameters (ξ_d, ξ_D), 2 location parameters (μ_d, μ_D), and the seasonal mean of extreme travel delay (Δ_j). The details are provided in Table 2.3.

Table 2.3. Algorithm for estimating the parameters of the CGEV distribution

Procedure *CGEVParametersEstimation*
Input: Historical travel time information for links in network G
Output: Parameters of the CGEV for each link
Initialization:
Assign an appropriate value to the time interval T
Assign an appropriate value to the block interval τ where $\tau \leq$ T
Step 1: Estimate parameters of CGEV related to period-to-period level and validate the model assumption

For each link $l \in G$

For each season *j=1, …, S*

For each time interval *t=1, …, M*

Set $D_{tj} \coloneqq \frac{1}{n_{tj}} \sum_{i=1}^{n_{tj}} d_{itj}$

Set $D_t \coloneqq D_{tj}$

For each block interval $i = 1, \ldots, n_{tj}$

Set $d_t \coloneqq d_{itj}$

Set $(\varepsilon_d)_i \coloneqq \frac{d_t}{D_t}'$

End for

Set $\varepsilon_{d\_t} \coloneqq \left\{(\varepsilon_d)_1, \ldots, (\varepsilon_d)_{n_{tj}}\right\}''$

End for

Set $\varepsilon_d \coloneqq \left\{\varepsilon_{d\_1}, \ldots, \varepsilon_{d\_M}\right\}$

End for

If Model 1

Determine parameters of GEV distribution (μ_d *and* ξ_d) by applying MLE method to the $\frac{N}{\tau}$ error term ε_d

End if

If Model 2

Set $\varepsilon_d \coloneqq \left\{\varepsilon_w, \varepsilon_p, \varepsilon_s, \varepsilon_f\right\}'''$ (Assigns block error terms in similar seasons across different years to the same set)

Determine parameters of GEV distribution (μ_d *and* ξ_d) by applying MLE method to the error terms in each season

End if

Set $\hat{\theta} \coloneqq \frac{1-\mu_d}{\Gamma(1-\xi_d)-1} \sqrt{(\Gamma(1-2\xi_d) - \Gamma(1-\xi_d)^2)}$

Set $\theta \coloneqq \frac{\sigma_{d_{tj}}}{\mu_{d_{tj}}}$

Compare how close is $\hat{\theta}$ to θ

Step 2: Estimate parameters of CGEV related to seasonal level and validate the model assumption

For each season *j=1, …, S*

Set $\mu_{D_j} \coloneqq \frac{1}{M} \sum_{t=1}^{M} D_{tj}$

Set $\Delta_j \coloneqq \mu_{D_j}$

For each time interval *t=1, …, M*

Set $(\varepsilon_D)_t \coloneqq \frac{D_t}{\Delta_j}$

End for

Set $\varepsilon_D \coloneqq \{(\varepsilon_D)_1, \ldots, (\varepsilon_D)_Q\}$

If Model 1

Determine parameters of GEV distribution (μ_D *and* ξ_D) by applying MLE method to the $\frac{N}{T}$ error term ε_D

End if

If Model 2

Set $\varepsilon_D \coloneqq \{\varepsilon_{D\_w}, \varepsilon_{D\_p}, \varepsilon_{D\_s}, \varepsilon_{D\_f}\}$ ‴ (Assigns time interval error terms in similar seasons across different years to the same set)

Determine parameters of GEV distribution (μ_D *and* ξ_D) by applying MLE method to the error terms in each season

End if

Set $\hat{\Theta} \coloneqq \frac{1-\mu_D}{\Gamma(1-\xi_D)-1}\sqrt{(\Gamma(1-2\xi_D)-\Gamma(1-\xi_D)^2)}$

Set $\Theta \coloneqq \frac{\sigma_{D_j}}{\mu_{D_j}}$

Compare how close is $\hat{\Theta}$ to Θ

End for

End for

′ $(\varepsilon_d)_i$ = maximum error term in block i

″ ε_{d_t} = Set of error terms in time interval t

‴ w: winter, p: spring, s: summer, f: fall

2.5.4 Performance of the CGEV distribution

In this section, the performance of the proposed CGEV distribution is evaluated in terms of the goodness of fit compared to the empirically observed data. The performance of the CGEV distribution is further compared with the performance of the traditional GEV distribution, taking into account both the goodness of fit and the complexity of the model. For the comparison, four

different models consisting of two CGEV models and two GEV models are constructed to be fit to the empirical data as follows:

a. Model 1: CGEV with different Δ_j for each season

In the first model, similar seasons are assumed to have similar Δ_j values, but the rest of the CGEV parameters are assumed to remain constant over different seasons, i.e., similar ξ_d, ξ_D, μ_d, μ_D. The choice of the constant and varying parameters in this model is based on the assumption that the PDF tail characteristics, i.e., related to the shape parameters ξ_d, ξ_D, and the translation of the PDF on the x-axis, i.e., related to the location parameters μ_d, μ_D, remain fairly constant over different seasons, while the magnitude of the average extreme travel delay observed on a link, represented by Δ_j, varies from one season to another (e.g., from spring to winter). It was further assumed that similar seasons across different years have similar Δ_j values (e.g., spring 2015 and spring 2016 have same Δ_j values).

Six years of travel time data, or twenty-four different seasons, results in four CGEV distributions to be fit to the empirical extreme travel delay data. These distributions share similar shape and location parameters, but they have a different value for Δ_j. Thus, 8 parameters consisting of 2 shape parameters, 2 location parameters, and 4 Δ_j, $j = 1,\ldots,4$, should be estimated to construct all the required CGEV models to describe the data. The Δ_j values in Model 1 are calculated according to Table 2.2, and the shape and location parameters are estimated based on the MLE method as described in Table 2.3.

b. Model 2: CGEV with different parameters over different seasons

While assuming similar shape and location parameters across different seasons, as in Model 1, simplifies the model and reduces the computational time in constructing the CGEV

distributions for different seasons as fewer parameters need to be estimated, seasonal variation in extreme travel delay data (e.g., from winter to spring) might lead to different tail and location characteristics of the PDF of the data. Thus, to investigate the plausible improvement in the goodness of fit, in Model 2, it is assumed that similar seasons across different years share similar shape and location parameters while the shape and location parameters vary across different seasons within a year.

Moreover, similar to what is assumed in Model 1, it is plausible that the extreme travel delay process, which is mainly a function of environment and weather conditions for similar seasons across different years, is fairly homogenous. This homogeneity can be modeled by allowing different Δ_j values for different seasons (e.g., from spring to winter), but common Δ_j values for similar seasons across different years which in turn generates 4 CGEV distributions to be fit to the empirical extreme travel delay data. In Model 2, 20 parameters, 8 shape parameters, 8 location parameters, and 4 $\Delta_j,\ j = 1,\ldots,4$, need to be estimated.

Model 2 increases the complexity of the model parameters estimation compared to Model 1 by relaxing the homogenous extreme travel delay process across different seasons of different years. The Δ_j values of the CGEV distribution are obtained as described in Table 2.2, and the shape and location parameters are estimated using the MLE method as illustrated in Table 2.3.

c. Model 3: GEV with different parameters for each season

Similar to Model 2, Model 3 relaxes the assumption that the tail characteristics and the translation of the PDF on the x-axis remain constant across different seasons. Moreover, similar to Model 2, it is assumed that the extreme travel delay process remains homogeneous across similar seasons of different years, and thus, similar seasons across different years share the same shape (ξ), location (μ), and scale parameters (σ). This assumption results in constructing 4 GEV models

(1 for each season) for which 12 parameters need to be estimated in total (4 shape parameters, 4 location parameters, and 4 scale parameters). Regarding the model complexity in terms of the number of model parameters, Model 3 is ranked between Model 1 and Model 2.

The 3 parameters of the GEV distribution in each season are obtained by applying the MLE method. Therefore, the GEV distribution is fitted to the $\frac{N}{4\tau}$ error terms ε_d, which exist in each season.

d. Model 4: GEV with same parameters across all seasons

In Model 4, all seasons across different years are assumed to share similar shape, location, and scale parameters. Model 4 is the simplest model discussed in this section as only three parameters need to be estimated for the whole database, regardless of the seasonality change in the observed extreme travel delay across a year.

The three parameters of the GEV distributions in Model 4 are obtained by applying the MLE method to the whole database, which requires fitting the GEV distribution to all $\frac{N}{\tau}$ error terms ε_d.

The sample results of the parameters of the CGEV distributions for Models 1 and 2 for the case of the 0.82 km section of Shaganappi Trail NW are shown in Table 2.4, following the models' specifications in Table 2.3. In addition, the CVs enhanced by the CGEV distribution (theoretical CVs) are compared to the CVs obtained from fitting a linear function to the seasonal empirical extreme travel delay data (actual CV) in Table 2.4 to validate the period-to-period and seasonal variation in extreme travel delay.

Table 2.4. CGEV distribution and model validation results for Models 1 and 2 for the case of the 0.82 km section of Shaganappi Trail NW

Model	Season	Constructed CGEV distribution							Slope		Ratio	
		Parameters					CV		θ	Θ	$\theta/\hat{\theta}$	$\Theta/\hat{\Theta}$
		ε_d	μ_d	ε_D	μ_D	Δ_j	$\hat{\theta}$	$\hat{\Theta}$				
Model 1	Spring	0.156	0.908	0.215	0.956	8.94	0.2018	0.0983	0.1257	0.1368	0.623	1.392
	Summer					8.98			0.1570	0.2075	0.778	2.111
	Fall					9.97			0.2680	0.0657	1.33	0.687
	Winter					8.84			0.1787	0.0409	0.885	0.416
Model 2	Spring	0.207	0.940	0.144	0.948	8.94	0.1343	0.1143	0.1257	0.1368	0.936	1.197
	Summer	0.030	0.926	0.109	0.901	8.98	0.1630	0.2149	0.1570	0.2075	0.963	0.966
	Fall	0.157	0.876	0.479	0.990	9.97	0.2720	0.0637	0.2680	0.0657	0.985	1.003
	Winter	0.168	0.916	0.303	0.979	8.84	0.1790	0.0519	0.1787	0.0409	0.998	0.788

Investigating the value of the ratios between the theoretical and actual CVs across different seasons for the examined link reveals significant deviations between the theoretical and actual CVs in Model 1, especially in the case of seasonal variability. To investigate whether this is the case at the network level, the monthly and seasonal ratios are evaluated for all the links in the Calgary road network. The results are shown in Figure 2.4.

As shown, in most cases, there is significant deviation between the actual and theoretical CVs across all the links in the network for Model 1. A possible source of the deviations is the generalization of the shape and location parameters across all seasons. Therefore, the assumption that the PDF characteristics remain constant over different seasons does not hold. The mean of the ratios is in the range of [0.915, 0.984] and [0.894, 1.098] for monthly and seasonal levels, respectively, which indicates that the extreme travel delay can be characterized by the CGEV distribution constructed in Model 2 for both monthly and seasonal variability. In addition, comparing the SD of the monthly and seasonal ratios across different seasons reveals more variations in the ratios at the seasonal level for both Models 1 and 2. The higher dispersion around

the mean ratio for the seasonal level compared to the monthly level is possibly due to the smaller sample size of the mean-SD points used to construct the mean-SD plots for the seasonal level (similar to what is depicted in Figure 2.2). Thus, these results are a relatively more biased estimation of the actual CV obtained from applying linear regression to the mean and SD of the extreme travel time as calculated for each season.

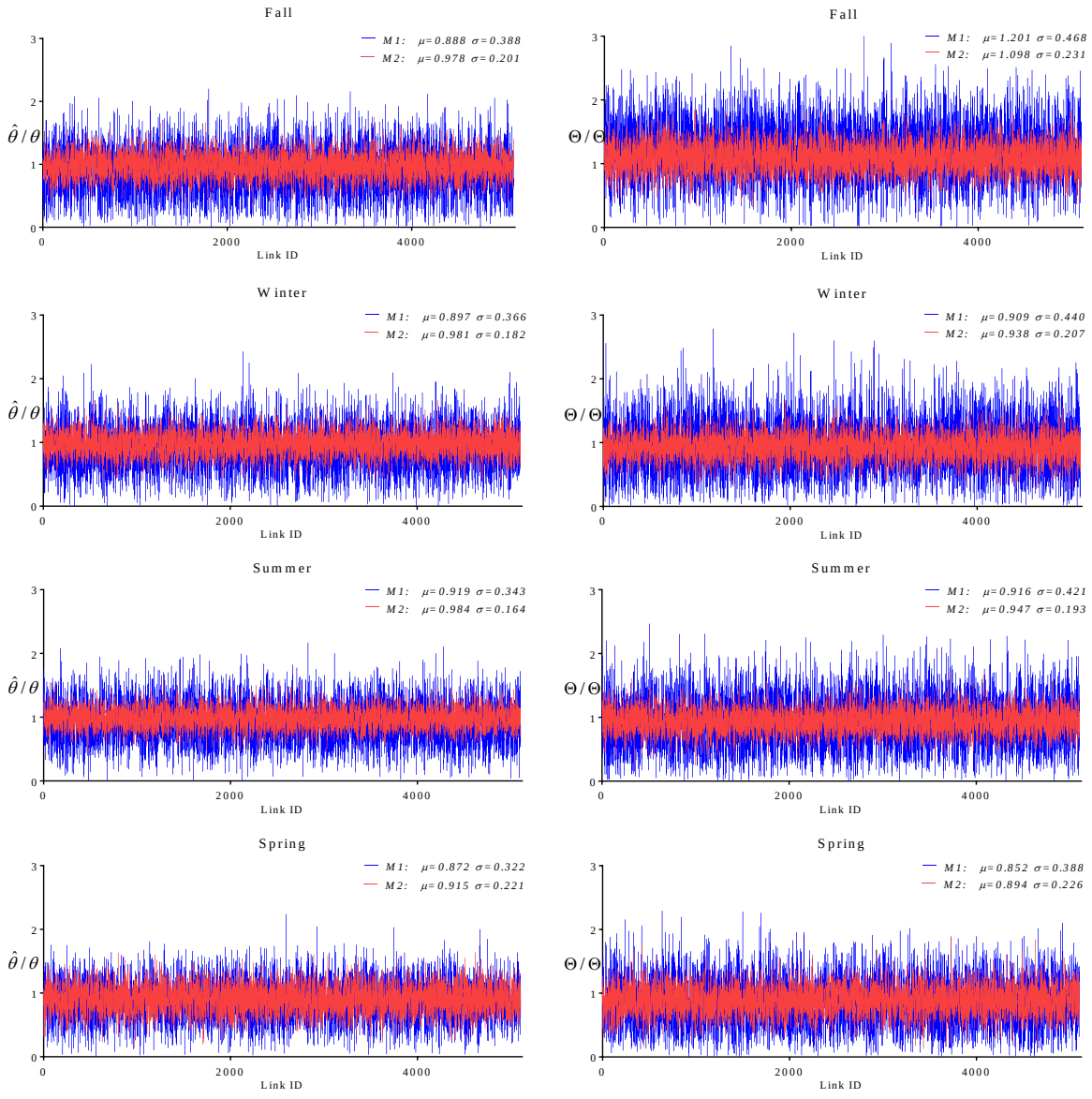

Figure 2.4. Monthly levels (left) and seasonal levels (right) ratios for the four seasons of the year calculated for all links with INRIX data in the Calgary road network

To compare the performance of the proposed CGEV models with that of the conventional GEV models and evaluate how well the distribution fits the empirical data, an appropriate measure of goodness of fit needs to be considered. The choice of this measure can be tricky as the number of the estimated parameters varies across different models. Since the number of parameters is not the same across different distributions, the maximum likelihood value is not a precise estimator of the goodness of fit because it only measures the goodness of fit without accounting for the complexity of the model (i.e., the number of estimated parameters). Since increasing the number of estimated parameters increases the goodness of fit and the complexity of the model at the same time, Akaike Information Criterion (*AIC*) (Akaike, 1998) is an ideal measure that accounts for both the goodness of fit and the complexity of the models. *AIC* is defined in Equation (2.31).

$$AIC = 2k - 2ln(L) \tag{2.31}$$

where *k* is the number of estimated parameters and *L* denotes the maximum value of the likelihood function. The model is better with a smaller *AIC*. While *AIC* rewards the goodness of fit, it penalizes the model for overfitting brought by a higher number of estimated parameters for each season. After obtaining the *AIC* for all constructed models, the *AIC* of each model is then subtracted from the *AIC* of the best candidate model in each season to obtain the relative *AIC* difference for each model across different seasons; $\Delta AIC_m = AIC_m - \min_N(AIC_n)$, *n =1,..,m,...N*. If the *AIC* difference in a season is low, the associated model is better in terms of fitting to the empirical data related to that season. $\Delta AIC_m = 0$ is set for the best-fitted model to the data.

The *AIC* is calculated based on the log-likelihoods obtained from fitting different models to the seasonal extreme travel delay data. For illustration purposes, the performance of different models in each season in terms of the log-likelihood value and across all seasons in terms of the

AIC value is compared in the case of the 0.82 km section of Shaganappi Trail NW N and is shown in Figure 2.5.

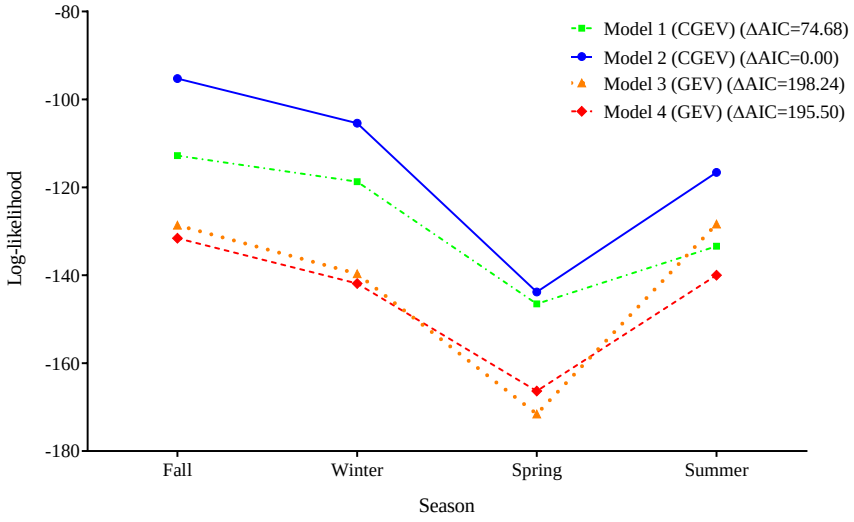

Figure 2.5. Values of log-likelihood and *AIC* under different CGEV and GEV distributions fitted to the extreme travel delay data

The relative log-likelihood value calculated over different seasons for Model 2 shows a superior fit with the empirical data followed by Model 1 then the other two GEV distributions. The performance of Model 4 (the least complex model) is not significantly far from Model 3 (the second most complex model), which necessitated comparing the *AIC* values to account for model complexity. The estimation of ΔAIC values indicates that the CGEV distribution represented in Model 2 is favored over the other three distributions across all seasons followed by Model 1, which has a smaller ΔAIC value compared to Model 3 and Model 4. However, while Model 3 provides a better fit with the empirical data (better log-likelihood value) compared to Model 4, the complexity of Model 3 leads to a higher value of ΔAIC and thus, Model 4 outperforms Model 3.

To evaluate the overall performance of the models and to see whether the same trend holds at the network scale, the ΔAIC values associated with each model are evaluated and compared for

every link in the Calgary road network. The results indicate that Model 2 has the highest log-likelihood value for every link in the network across different seasons, which indicates a superior fit with the empirical data. Taking the model complexity into account, Model 2 has the lowest ΔAIC in 4995 out of 5104 links in the network, and thus, it performs better compared to the other three distributions. Moreover, for the 117 links in the network for which Model 2 is not the best fitted distribution to the empirical data, Model 2 is ranked either second (101 times) or third (8 times) with no record of Model 2 being ranked fourth in any of the links; consequently, Model 2 outperforms the other distributions. The superiority of Model 2 comes from the fact that Model 2 accounts for all seasonality aspects of the extreme travel delay data by allowing different parameters to be estimated across different seasons, i.e., different shape and location parameters and different Δ_j values for different seasons.

The CGEV distribution represented in Model 1 is the second-best model followed closely by Model 3. While Model 3 outperforms Model 1 in most of the links when the log-likelihood values are compared (due to the generalization of the shape and location parameters across different seasons in Model 1 as discussed before), Model 3 has many more parameters to be estimated in each season, which leads to a more complex model and, thus, a higher ΔAIC value. In other words, while the GEV distribution represented in Model 3 is heavily penalized because of its extra parameters (12 parameters to be estimated compared to 8 parameters in Model 1), the actual fit tends to be better in Model 3 compared to Model 1 without factoring in the complexity. Model 3 over-generalizes the characteristics of extreme travel delay by ignoring monthly and seasonal variation in the extreme travel delay, which results in a poor fit to the empirical data. This over-generalization is not observed in Model 1 as Model 1 accounts for seasonality in the data by allowing different Δ_j values to be estimated across different seasons, which leads to its superior

performance compared to Model 3. Model 4, the least complex model, is the least favored model compared to the other three distributions and provides a relatively poor fit to the empirical data; it is ranked as the worst distribution in 4704 of the links in the Calgary road network

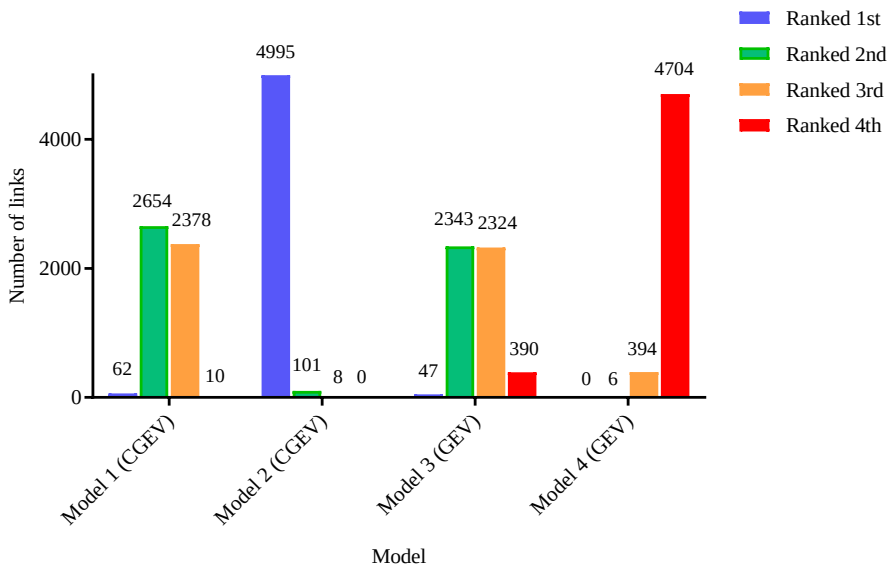

Figure 2.6. Model ranking and the associated frequency based on ΔAIC values across all links in the Calgary road network

Figure 2.7 shows an example of detailed histograms of extreme travel delay across different seasons, presented along with the density curves associated with the constructed models for the case of the 0.82 km section of Shaganappi Trail NW N, Calgary. As shown, the CGEV family consisting of Model 1 and 2 better represents the observed underlying extreme travel delay data across different seasons compared to the GEV family (Model 3 and 4). The only exception is the summer season, in which Model 3 from the GEV family provides a closer fit to the observed frequency of data compared to Model 1. This finding is consistent with the log-likelihood values reported in Figure 2.5 for Model 1 and Model 3 for the summer season. In addition, Figure 2.7 indicates that Model 2 is capable of modeling a wide range of extreme traffic conditions given the

four unique shape and location parameters for each season (ε_d, ε_D, μ_d, μ_D); it characterizes the monthly and seasonal variation in data and adapts to the temporal evolution of the extreme travel delay distribution through the season-sensitive varying mean parameters (Δ_j, $j = 1,\ldots,4$). While flexibility usually comes at a cost of having a more complex model, the CGEV distribution constructed in Model 2 does not add much to the model complexity as it requires estimating additional mean parameters that are fairly easy and straightforward to calculate across different seasons. Further, Model 2 only has one more season-sensitive parameter for each season compared to Model 3. This level of flexibility is not offered by traditional GEV distributions in Models 3 and 4 as they fail to account for the extreme travel delay distribution given the large variance in extreme travel delay distributions over different time periods for which the GEV parameters are calibrated. Therefore, in the context of road network vulnerability, CGEV distributions can offer a unifying framework in assessing, comparing, and describing the extreme travel time distribution considering the seasonality in the data.

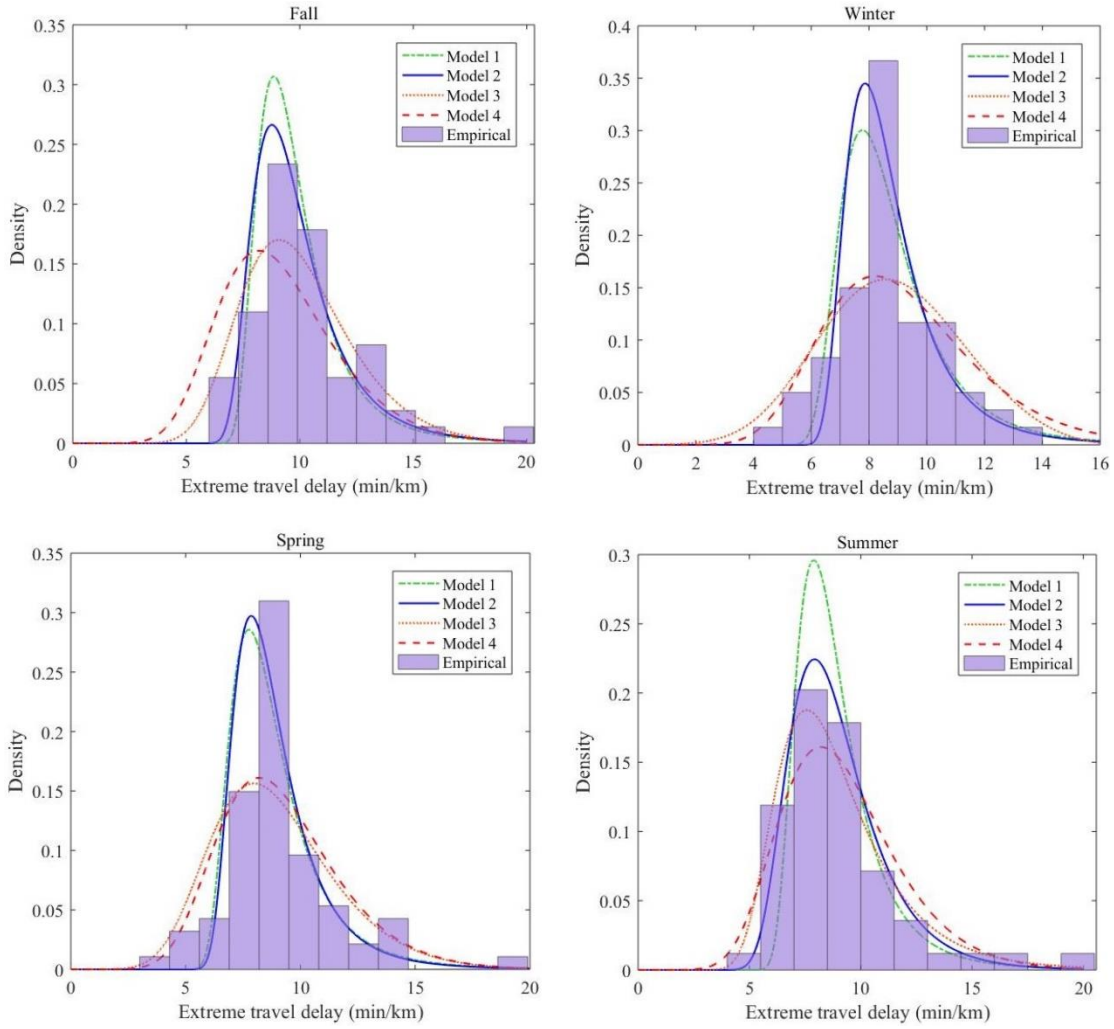

Figure 2.7. Extreme travel delay distributions and the fitted density curves for different seasons for the 0.82 km of Shaganappi Trail NW N, Calgary

I will be using this interesting feature of the CGEV distribution to obtain a reasonable prediction of the return level associated with the extreme travel delay distribution constructed for each season. The return level is then used to build a data-driven measure of vulnerability to find the vulnerable elements of the Calgary road network.

2.5.5 Return level and further model validation

The concept of return level has been used extensively in the literature to account for the impact of extreme events in different fields, including environmentally related subjects such as predicting maximum water level (Mudersbach and Jensen, 2010), predicting maximum daily rainfall (Coles and Pericchi, 2003), and predicting maximum global temperature (Cheng et al., 2014), and subjects related to financial markets, to measure the value at risk (Gilli, 2006). In the context of transportation, the concept of return level has been employed widely in predicting the severity of car crashes (Åsljung et al., 2017; Åsljung et al., 2016; Tarko and Songchitruksa, 2005; Zheng et al., 2014), measuring the exposer of road networks to extreme weather conditions (Schlögl and Laaha, 2017), and predicting the severity of aviation incidents (Panagiotakopoulos et al., 2009).

The measures of return level, i.e., extreme quantiles of a block maximum distribution, and return period, i.e., the period between two succeeding exceedances of the corresponding return level value, provide important information for decision-making processes and assessing the impact of extreme events in terms of the magnitude of the observed extreme travel delay. If $F(d_t)$ is the CDF of the maximum observed extreme travel delay over successive non-overlapping time blocks of equal length and $d_p(l)$ is the return level of link l associated with the return period $\frac{1}{p}$, to a reasonable degree of accuracy, then the level $d_p(l)$ is expected to be exceeded on average once every $\frac{1}{p}$ period of time. More specifically, $d_p(l)$ is exceeded by the block maximum in any block with the probability of p; therefore, $F\big(d_p(l)\big) = 1 - p$. In the case of the GEV distribution, the return level of link l, i.e., $d_p(l)$, is calculated as follows:

$$d_p(l) = \mu_l - \frac{\sigma_l}{\xi_l}\left[1 - \{log(1-p)\}^{-\xi_l}\right] \tag{2.32}$$

where μ_l, σ_l, and ξ_l are the location, scale, and shape parameters of the GEV distribution associated with link l.

Calculating the return level of the CGEV distribution is not as straightforward as in the case of the GEV distribution. No closed-form equation exists for the return level of the CGEV distribution; however, the return level can be calculated for the desired period of time numerically. As the return level is the quantile of the CDF for the return period $\frac{1}{p}$, the return level can be obtained by trial and error where different values of d_t are evaluated in the CDF function successively until the CDF function converges to $1 - p$ as shown in Equation (2.33).

$$F\big(d_p(l)\big) = \int_{-\infty}^{d_p(l)} f(d_t)\, dd_t = 1 - p \tag{2.33}$$

To evaluate how accurately the observed return level can be predicted by the analytical return level obtained from the CGEV distribution, the entire 6 years travel time database associated with each link in the Calgary road network is split into two 3-year travel time databases (one from December 2013 to December 2016 and the other from January 2017 to December 2019); the analytical return level is calculated for the first database and is validated by comparing it to the observed mean return level estimated from the second database. The observed mean return level in each season (e.g., fall) is calculated by choosing the maximum extreme travel delay in each of the three seasons in the second database (i.e., fall 2017, fall 2018, and fall 2019) and calculating the mean of these three extreme observations. To further validate the fit of the CGEV distribution with the empirical

data compared to the traditional GEV distribution, seasonal return level ($\frac{1}{p}$ = 12) is calculated for both Model 2 and Model 3 and compared with the seasonal observed mean return level.

Table 2.5 shows the return levels of the extreme travel delay associated with Model 2 and Model 3 across different seasons calculated for the 0.82 km section of Shaganappi Trail NW N, Calgary (Link 1) and a 1.9 km section of Deerfoot Trail between Country Hills Blvd NE and Airport Trail (Link 2). The closest predictions to the actual return level and the corresponding observed mean return level are highlighted in bold.

Table 2.5. Sample comparison of the return levels obtained from the CGEV and GEV distributions for 2 links in Calgary

Link		Spring	Summer	Fall	Winter
1	Return level (Model 2)	**11.41**	**12.34**	**12.52**	**10.93**
	Return level (Model 3)	15.93	11.95	15.02	14.23
	Observed mean return level	12.68	13.57	13.31	11.58
2	Return level (Model 2)	**5.78**	**6.02**	**9.81**	**7.03**
	Return level (Model 3)	5.48	4.87	6.76	5.32
	Observed mean return level	6.01	6.76	10.72	7.49

As shown, the CGEV distribution in Model 2 outperforms the GEV distribution in Model 3 because it provides a more accurate prediction of the actual mean return level across different seasons for both case studies. To validate whether the same statement holds for other links, the analytical return levels for both Model 2 and 3 are obtained for all links in Calgary road network across different seasons and are compared to the actual mean return levels. The percentage error of prediction is depicted as a heatmap across different seasons in Figure 2.9.

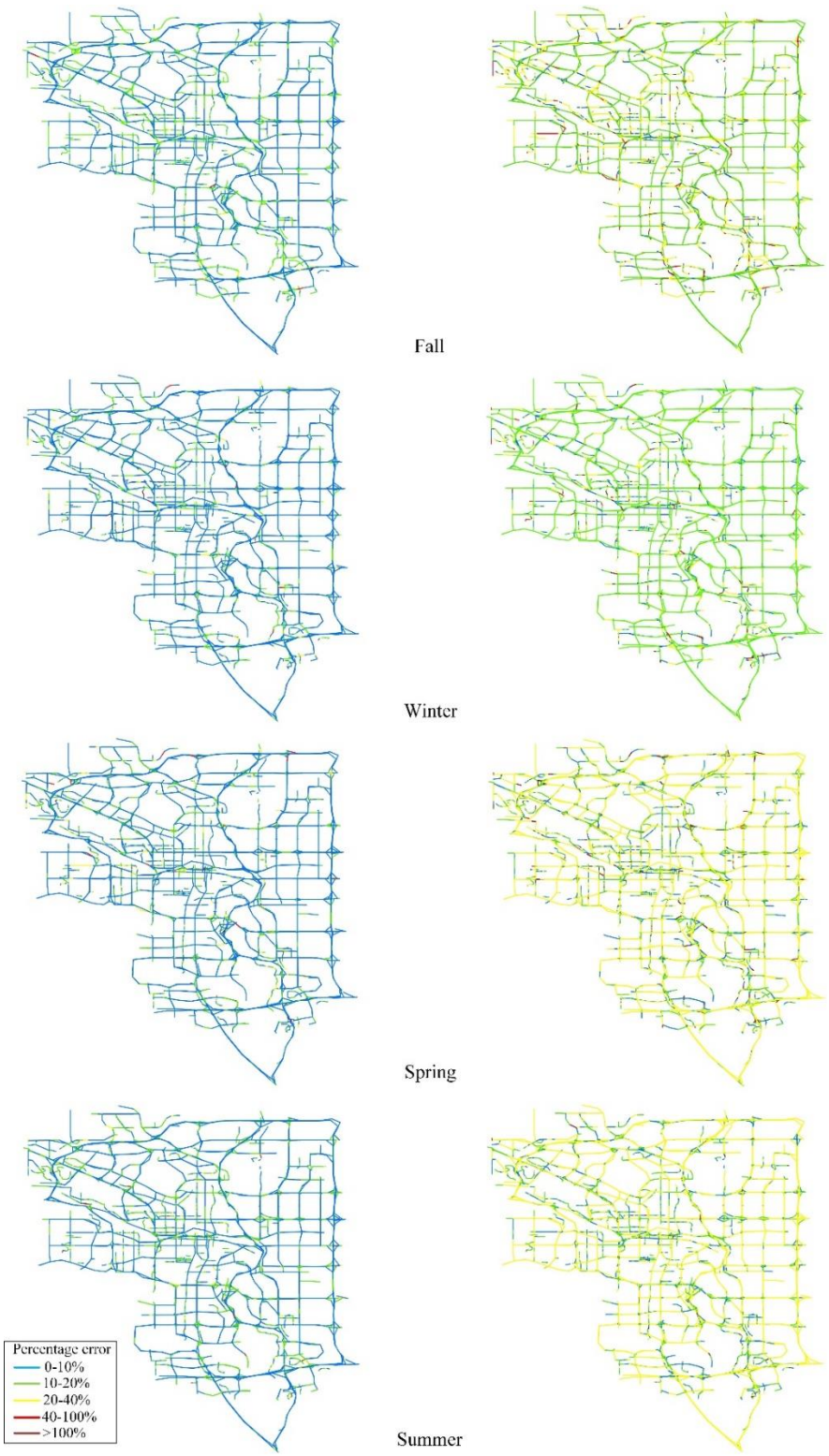

Figure 2.8. Percentage error of estimation across different seasons for Right: Model 3 and Left: Model 2

As shown in Figure 2.8, Model 2 consistently outperforms Model 3, especially for the seasons of spring and summer because the errors of estimation are significantly larger in Model 3. While the CGEV model performance remains fairly consistent over different seasons, the GEV distribution performance is much worse than that of the CGEV distribution for the seasons of spring and summer. This superior performance of the CGEV distribution further supports the merit of the CGEV distribution in modeling the variability in a large size network and minimizing the error in predicting the actual return level.

2.6 Applications of CGEV distribution in road network vulnerability analysis

Determining the critical links in a road network is essential for resilience analysis because it contributes to improved resource management by proactively preventing events that contribute to rare events in a network (e.g., collisions). The frequency of extreme events occurring during a certain time period can be estimated by analyzing travel time data. The travel time of a link is likely high in the presence of rare events. As the probability of rare events is very low, I expect that high travel times occur less frequently. Therefore, considering the distribution of the travel time data for a given link in a transportation network, greater travel times have a lower probability and, thus, are located at the tail of the distribution. EVT is a powerful tool to estimate the frequency of extreme travel delay occurrence. The probability of each link closure can be modeled alternatively by measuring the return level of extreme travel delays associated with each link in a road network. One advantage of using return level as a proxy of the frequency of exceedance is that the return level accounts for the magnitude of an extreme travel delay that occurs on average once every $\frac{1}{p}$ period. In other words, it models the magnitude of the extreme travel delays and takes into account the stochastic nature of extreme travel delays as the return level is directly related to

the estimated parameters of the GEV distribution as shown in the return level measures indicated (i.e. Equations (2.32) and (2.33)). It is desirable to use the normalized return level, i.e., return level pace, as a performance metric instead of the more traditional measure of vehicle count since the goal is to measure the impact of traffic congestion during extreme events (Donovan and Work, 2017).

While return level addresses the magnitude and stochasticity associated with the occurrence of rare events in a given link, it does not model how important a link is in a given road network compared to other links. For instance, a minor link with a low daily travel demand can have a larger return level of extreme travel delay compared to a congested major link with a high daily demand for travel. To account for the link importance, the importance of each link in a given network is quantified by estimating the weights of each link contributing to network performance. Adopting the idea of trip importance introduced by Zhong et al. (2019), a new link importance measure is introduced in this study. Given L links, denote $X_k^i(l) = \frac{u_k^i(l)}{w(l)}$ as the pace distribution of link $l = 1, \ldots, L$ in the time interval k of day i and $R_k^i(G) = \frac{\sum_{l=1}^{L} w(l) X_k^i(l)}{\sum_{l=1}^{L} w(l)}$ as the network-scale pace distribution of network G at time interval k of day i, where $u_k^i(l)$ is the travel time of link l in the time interval k of day i, and $w(l)$ is the length of link l; the following optimization problem is solved to determine the importance of each link on network performance:

$$Z = min \sum_{i=1}^{N} \sum_{k=1}^{T} \left(E\left[R_k^i(G) - \sum_{l=1}^{L} a(l).X_k^i(l) \right] \right)^2 \quad (2.34)$$

s.t. $\sum_{l=1}^{L} a(l) = 1$

$0 \leq a(l) \leq 1$

Where $a(l)$ is the importance weighting of link l. The results obtained from solving the above equation is a number between 0 and 1 for each link in the network. As $a(l)$ increases, the more link l contributes to the network performance; thus, links with the highest values of $a(l)$ have the most substantial contribution to network performance.

A new vulnerability index is then introduced by considering both the stochastic nature of extreme events, expressed by the return level of the extreme travel delay, and the link importance:

$$VUL_l^G = a(l).d_p(l) \tag{2.35}$$

where VUL_l^G is the vulnerability index of link l in network G and represents the relative impact of the closure of link l on network performance. Based on the obtained vulnerability index, the vulnerable ranking of link l, denoted by $rank_l^G$, can be determined in network G.

Table 2.6. Algorithm for ranking vulnerable links

Procedure *FindingVulnerableLinks*
Input: Historical travel time information for links in network G
Output: Network vulnerability ranking for each link
Step 1: For each link $l \in G$ Fit the CGEV distribution by procedure *CGEVParametersEstimation* Calculate the return level $d_p(l)$ using Equation (2.33) Calculate the importance weighting of link l among all the other links in network G using Equation (2.34) Calculate the network vulnerability index VUL_l^G using Equation (2.35) End for
Step 2: Determine the $rank_l^G$ values by sorting the VUL_l^G values in a descendent order

For the analysis, the majority of roadways in Calgary, AB, Canada, (around 85% of roadways, consisting of 5104 links) were selected. The selected corridors include provincial

highways and major road networks including freeways, expressways, arterials, and local and minor roads within the metro Calgary boundary. Figure 2.9 shows the critical links of the Calgary road network.

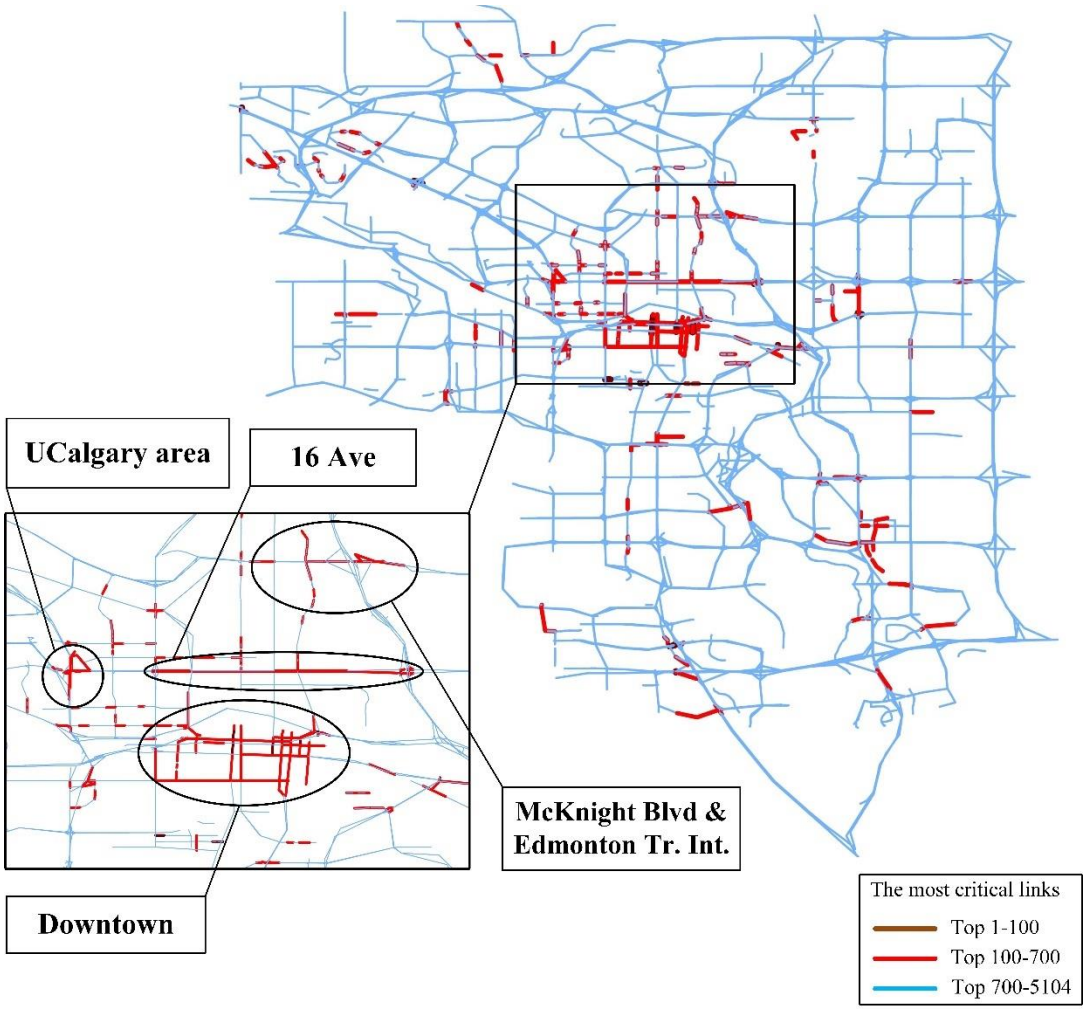

Figure 2.9. The critical links in the Calgary road network

Notable areas consisting of clusters of critical links are highlighted in Figure 2.9. The most critical links are the roads within the Calgary downtown area. Calgary downtown is the second-largest concentration of head offices in Canada and arguably the densest downtown area of any city of its size in North America (Canada West Foundation, 2010). Due to the high concentration

of businesses and high population density, Calgary downtown has a very high demand for daily travel. The high travel demand to downtown Calgary might be one of the most important factors contributing to the vulnerability of these links. Despite being a dense urban area (measured by the number of links per unit area), most of the critical links in this area (4th, 5th, 6th, and 9th Ave, and 4th, 5th, and 9th St SW) are one-way roads with limited alternative access roads to these links, i.e., low connectivity. In most cases, the alternative roads to these links are other links that are two blocks away. In addition, to access these one-way roads, drivers need to use other one-way roads. Therefore, because of the low connectivity between one-way streets, high network density in the downtown area cannot easily mitigate the consequences of road closure within the area. Thus, the lack of alternative links to these roads could be another influential factor in the vulnerability of these links. One last influential factor is the lack of cyclicity in downtown Calgary. Due to the high number of one-way streets, a few numbers of random walk lead to a cycle back to a previously visited node in the downtown area. Lack of cyclicity further affects the capability of the downtown area to recover from the consequences of link failure. This finding is consistent with the finding in Zhang et al. (2015), who reported cyclicity to be positively correlated with resilience.

Another city area consisting of a considerable number of critical links is in the vicinity of the University of Calgary (the largest university in Southern Alberta) and major roads connecting other city areas to the university, i.e., 16 Ave and part of Crowchild Trail. Banff Trail and part of Crowchild Trail in the vicinity of the university are links with high traffic demand and a very high annual collision rate (around 100 incidents per year). In addition, Banff Trail is the only link that connects 16 Ave and Crowchild Trail, two major expressways. This roadway design makes Banff trail highly vulnerable to disruptive events.

16 Ave is a major urban arterial road in the City of Calgary that forms 26 km of the Trans Canada Highway and connects Calgary to Banff and Medicine Hat. 16 Ave crosses major expressways (Sarcee Trail, Crowchild Trail, Deerfoot Trail, and Stoney Trail), but only the part between Deerfoot Trail and 14th St has extensive commercial development. In addition, this part of 16 Ave connects numerous demand generators in the City of Calgary to the Southern Alberta Institute of Technology (SAIT) and the University of Calgary. While this part of 16 Ave is a major link with high travel demand, the network density is not high in the vicinity of this link, especially compared to the Calgary downtown area. Furthermore, most of the links in the Calgary network are radial links connecting the center with very few east-west connections. Thus, most of the east-west connections load onto 16 Ave, which makes the travel demand excessive on this link. Moreover, the high numbers of signalized intersections with complex signal timing and high collision rate (50 incidents per year) compared to similar expressways in the Calgary road network (Farhan, 2019) make this part of the road more vulnerable to disruptive events.

Comparing my findings in the Calgary downtown and 16 Ave/Banff Trail cases with the findings of Chen et al. (2012), my results are consistent and it can be concluded that for congested urban areas, those roadways that have limited alternative access links tend to be vulnerable to disruptive events.

Another network area consisting of highly vulnerable links is the intersection of McKnight Blvd and Edmonton Trail. This part of the network is a low network density area with high travel demand, which directs the traffic demand from/to the NE area and Deerfoot trail from/to Calgary downtown via Edmonton trail. One possible reason contributing to the vulnerability of this area is the relatively high collision rate due to the sudden decrease in the speed limit on McKnight Blvd from 70 km/h to 50 km/h in the vicinity of the McKnight Blvd/Edmonton Trail intersection.

Another reason is the lack of alternative links to this area, which makes the area not robust in terms of mitigating the consequences of network disruption. Comparing the case of McKnight Blvd/Edmonton Trail with the previously discussed network areas, it can be inferred that for low-density network areas, the demand level plays a much more important contributing factor to link vulnerability.

CHAPTER 3: ROAD NETWORK VULNERABILITY ANALYSIS CONSIDERING PROBABILITY AND CONSEQUENCES OF DISRUPTIVE EVENTS

3.1 Introduction

3.1.1 Background and motivation

A robust and resilient transportation network is the backbone of an advanced society. A robust transportation network facilitates the movement of passengers, goods and services, and contributes to a flourishing economy. A transportation network, however, can be vulnerable to certain disruptive events that range from regular fluctuations and recurrent events to irregular and non-recurrent events. Recurrent events happen periodically in a road network and are the outcomes of the within-day variation of travel demand. Non-recurrent events, on the other hand, result from extreme and disruptive occurrences leading to severe deterioration of network performance. The possible sources of such unexpected events are natural (e.g., disasters, extreme weather condition), artificial (e.g., traffic collisions, major road repairs, social events) or technical (e.g., signal failure at a major intersection) (Li, 2008). Other classification of the causes of disruptive events can be further categorised as internal, external and intentional incidents interference (Mattsson and Jenelius, 2015). Internal events are those originating from the system and are caused by mistakes made by transportation authorities and staff (e.g., technical failure) and external events are mostly related to natural phenomena (e.g., extreme weather conditions) or artificial events caused by intentional interferences (e.g., terrorist or cyber attacks).

Non-recurrent events can result in long delays, spillbacks and network gridlocks. While the probability of occurrences of disruptive events are very low, the impact of these events on the performance of the transportation network can be huge. Therefore, in the transportation network, it is important to identify the vulnerable elements affected by such disruptive event. Such analysis

helps to manage the potential risks resulting from these events and to better alleviate the disruptions to improve all aspects of transportation network performance. Vulnerable network elements (e.g., links, nodes) are part of a network responsible for a sharp decrease in traffic operation performance caused by capacity restrictions due to an incident or special event (Yperman & Tampere, 2006). Searching for vulnerable elements of a road network aims to identify potential critical elements in a transportation network. Once these elements are identified, network robustness can be improved by either reinforcing these elements or by improving alternative parallel routes (Matisziw and Murray, 2009).

Numerous evaluation approaches and techniques have been considered in the literature to identify the critical road infrastructure. The evaluation approach is strongly linked to how a transportation network is represented for the purpose of the analysis. Depending on whether a transportation network is represented as an abstract graph or not, different measures of vulnerability are used in the literature and are discussed in Section 3.1.2.1. Apart from the evaluation approach, depending on how the scale of a disruptive event and its associated impact area is measured in a transportation network, different vulnerability analysis techniques were employed and are discussed in Section 3.1.2.2.

While previously discussed techniques and approaches have been employed to measure the vulnerability of network individual components (i.e., nodes and links) with the presence of disruptive events, these techniques deal with identifying the critical links of a transportation network, measuring only the consequences of the link closure with little consideration given to the probability of link closure. The probability of link closure or failure is important as some of the links in a transportation network are more susceptible to disruptive events than others. Instances of such links are links with problematic geometry design, major links with higher fluctuations in

travel demand and/or major highways with high merging, diverging and weaving traffic activity, which makes such links more prone than others to bottleneck formations and thus traffic breakdowns and possibly collisions. Moreover, depending on their type, disruptive events have a different probability of occurrence and their impacts differ case by case. The need for a more comprehensive approach was underlined in several previous studies. Berdiaca (2002), Chen et. al (2007), Erath et al. (2009), Tampere et al. (2007) and Watling and Balijepalli (2012) specified that vulnerability analysis should be regarded as risk analysis. Defining transportation system vulnerability as society's risk of transportation system disruptions and degradations, Jenelius (2009), Jenelius and Mattsson (2015) and Mattsson and Jenelius (2015) further specified that a vulnerability scenario should be formalised as a ''triplet'': 1) a description, 2) the probability and 3) the consequences (measure of damage) of that scenario. Therefore, more comprehensive vulnerability analysis is needed to simultaneously incorporate the probability and the consequences of disruptive events.

Moreover, important factors that affect the vulnerability of road links (e.g., poor geometry design) are usually left unnoticed by the analytical vulnerability approaches discussed in the literature. Multi-year observed travel time data carries important information that also reflects the historical spatial and temporal occurrence of link failure and their impacts. In this research, copious multi-year travel time data is further explored to identify the occurrence of disruptive events, examine the probability of occurrence of resulting extreme travel delays and evaluate their impacts on overall network performance. As a case study, this chapter uses the City of Calgary where rich historical travel time and incident data are available. The theoretical contribution of this research is the development of a novel impact area vulnerability analysis approach that incorporates both the probability and consequences of disruptive events to find the most vulnerable links in a road

network. In addition, this research contributes to the body of knowledge methodologically by proposing a dynamic statistical approach to determine the spatiotemporal impact of disruptive events. In addition, the obtained spatiotemporal impact is used to determine the probability of the impact resulting from occurrences of disruptive events. A risk-based vulnerability index is also developed using the concept of limit state function that is borrowed from structural reliability analysis.

3.1.2 Review of the previous studies

3.1.2.1 Evaluation approaches of vulnerability analysis

Road network vulnerability analysis has been recently receiving much attention in literature. Depending on how a road network is represented in the vulnerability analysis problem, two distinct evaluation approaches exist with little overlap: 1) topological and 2) system-based vulnerability analysis of road networks.

In the topological vulnerability analysis approach, a road network is typically represented as an abstract network (graph) consisting of nodes and links. Links can be represented as either directed or undirected links. Depending on the analysis application, links can also be considered unweighted (i.e., all links have similar lengths) or weighted (i.e., links have different lengths). Topological vulnerability studies usually deploy connectivity and network efficiency indices and other topological properties such as degree centrality and betweenness centrality as measures of vulnerability to assess the socioeconomic consequences of link closure (Demšar et al., 2008; Kurauchi et al., 2009; Duan and Lu, 2014). Denoting a connected network by *G*, the network efficiency indices usually represent the average distance between all possible pair of nodes in network *G* in the case of weighted links or the average number of links forming a route, in the case

of unweighted links. Various methods and techniques were employed in the literature to determine the vulnerable elements in road networks. As an example of recent topological network vulnerability studies on road transportation network, Bell et al. (2017) employed spectral analysis to find critical network cuts (i.e., cuts with least capacity normalized by the relative size of the network in either side of the cut) and thus the potential flow bottlenecks in road networks. Computational efficiency is one of the remarkable characteristics of the spectral analysis approach, which neither needs the origin-destination (OD) demand information nor the path-flow enumeration. Gao et al. (2019) conducted a topological vulnerability analysis considering both the quantitative aspect of connections between nodes (e.g., number of nodes connected to a specific node) and the information on the interaction level between nodes (e.g., how easily traffic can transfer from one node to its neighboring node). Assuming undirected and unweighted links, they make use of probability measures to describe the traffic flow between nodes. López et al. (2017) examined the impact of network topological configuration on network vulnerability and identified the vulnerable links by measuring the change in the traffic flow distribution associated with each link removal in the network. Other important examples of recent topological network vulnerability studies include important analysis conducted on public transport networks (Li et al., 2019; Ye and Kim, 2019) and air transport networks (Janić, 2015; Klophaus and Lordan, 2018; Lordan et al., 2014).

On the other hand, system-based vulnerability studies similarly employ graph representation of real networks consisting of weighted links with weights corresponding to actual link length, travel cost, etc. To identify the critical elements, transport network performances are examined through studying the interaction between travel demand and supply in terms of the increase in the total travel cost of the network. Several techniques and indices have been introduced

in system-based vulnerability literature to determine the vulnerability ranking of the examined network *G*. Almotahari and Yazici (2019) introduced a new vulnerability index taking into account traffic flow, link travel time and marginal travel cost associated with each link in the network. They also account for the importance of each link by assigning a weight based on the proportion of the flow that each link carries as compared to the total network flow. García-Palomares et al. (2018) conducted a vulnerability analysis on a road network in Spain. They introduced a vulnerability index based on three different accessibility measures including weighted average travel time, potential accessibility capacity, and daily accessibility. The accessibility measures used in this study mainly consist of travel time between each pair of centroid cities and the centroid population. Other interesting examples of system-based vulnerability analysis include those on public transport networks (Jiang et al., 2018; Lu, 2018; Yap et al., 2018) and air transport network (Darayi et al., 2017; Voltes-Dorta et al., 2017).

The fact that topological vulnerability approaches need limited data regarding the O-D demand and individual link traffic makes this approach computationally realistic and relatively easier to perform compared to system-based vulnerability approaches. The topological approach is also a more attractive choice of vulnerability analysis for networks without detailed network information. However, this approach does not totally reflect a realistic representation of the behavioural responses to disruption. Moreover, topological approaches do not take into consideration the dynamic effects of disruption, such as traffic congestion from disruption in any given link resulting in a significant increase in travel time on the neighboring links. System-based vulnerability approaches overcome some of these limitations as the travellers' response to disruption can be reasonably predicted. While the discussed analytical vulnerability approaches provide invaluable insights in understanding the importance of each link and the interactions

between nodes in a given network *G*, they usually overlook important contributing factors that affect the vulnerability of road links such as extreme weather condition, link specific geometry design, poor downstream traffic signal timing and/or links that are prone to higher frequency of collisions. In this study I make use of observed historical network wide link travel time and incident data which carry this wealth of information in order to derive a data-driven vulnerability index.

3.1.2.2 Employed techniques in vulnerability analysis

In any given transportation network, the critical links can be identified using different methods. The most used approach is to identify the critical links using a full network scan (Jenelius et al., 2006; Chen et. al., 2007; Du et al., 2014). In this approach, all the links in the transportation network are removed iteratively and the resulting impact is measured in terms of the increase in the generalized travel cost (e.g., increase in the travel time) of the passengers travelling all over the network. The vulnerability ranking is then determined to indicate the sorted order of the links in the transportation network ranging from the most to the least vulnerable. While the full network scan approach assures obtaining the global vulnerability ranking, it can be computationally intensive as the impact of link closure in terms of the reduced network performance needs to be evaluated for all links. The task of identifying the critical links may become computationally infeasible for a large-scale transportation network such as the Chicago regional network which has 39,018 links (Dial, 2006), especially in the case of system-based vulnerability approach.

To overcome the intensive computational time required to measure the impact of removing all the links, a pre-selected links approach based on certain strategies is used in the literature (D'Este and Taylor, 2003; Knoop et al., 2012; Taylor and D'Este, 2004; Taylor and Susilawati,

2012). In this approach, the analysis is performed on only these set of pre-selected candidate links, which correspond to high path choice probability between origin-destination nodes (Duan and Lu, 2014; Jenelius and Mattsson, 2015; Demšar et al., 2008). While this approach is much faster than the full network scan approach, obtaining global vulnerability ranking is not guaranteed. Moreover, the choice of pre-selected links is sensitive to the underlying strategy of choosing the candidate links.

In view of the above, an impact area vulnerability analysis approach was first introduced by Erath et al. (2009) and then Chen et al. (2012) to identify the most critical links in a congested large-scale transportation network. In this approach, the consequences of a link closure are evaluated within its local impact area and not within the whole network. The impact area vulnerability approach is based on the empirical findings that the impact of a link closure is mainly limited to the adjacent links and nodes and does not disperse throughout the whole network. As the local impact area is relatively small, the computational cost required for assessing the consequences of all possible link closures within the impact area, are moderate. Chen et al. (2012) proposed an impact area vulnerability approach to find the critical links in the Hong Kong road network. They considered the impact area associated with each link closure in the network deterministic and similar across different links in the network. One of the limitations of the analytical impact-area approaches discussed in the literature is that the size of the impact area cannot be adequately modeled. In fact, the size of the impact area not only varies from link to link but is also a by-product of the various factors which affect the O-D demand, including the within day and day-to-day fluctuation of travel demand, as well as weather conditions. These factors contribute to how severely neighbouring links are affected by disruptive events. In this study, an impact-area approach is taken into consideration and the spatial and temporal impacts of historical

disruptive events are measured using various data-driven and machine learning techniques to obtain the vulnerability ranking. While disruptive events can be categorized as planned and unplanned events, the focus of this study is on the impact of unplanned disruptive events. In fact, the behavioural demand response to planned disruptive events (e.g., work zones and other road works) is totally different from one of the unplanned disruptive events (e.g., collision) in terms of the demand elasticity with respect to learning and rerouting, as well as possible mode choice and destination choice shift. While the response to planned events is gradual and demand can be assumed relatively inelastic during these events, the demand response to unplanned events is highly elastic with respect to detouring (Erath et al., 2009). This makes analyzing the impact of planned events more challenging. Quantifying the impact of planned events is also more difficult to capture and necessitates evaluating the change in day-to-day travel demand pattern. Table 3.1 summarizes some of the findings in recent road network vulnerability analysis.

Table 3.1. Classification of recent road network vulnerability studies by evacuation approach and employed techniques

Study	Evaluation approach	Employed technique	Note
López et al. (2017)	Analytical/topological	full network scan	(i) Uses flow autocorrelation for more realistic representation of vulnerability (ii) No realistic representation of the behavioural responses to disruption (no traffic assignment)
Duan and Lu (2014)	Analytical/topological	pre-selected links	(i) Evaluated network vulnerability at three different granularities; segment-based, stroke-based and community-based. (ii) Not very realistic representation of flow distribution during disruptive events (iii) No consideration of the behavioural responses to disruption
Bell et al. (2017)	Analytical/topological	pre-selected links	(i) Uses spectral partitioning to identify flow bottlenecks in networks (ii) Computationally realistic and easy to perform (iii) No realistic representation of the behavioural responses to disruption
Gao et al. (2019)	Analytical/topological	pre-selected links*	(i) Characterizes evolution of vulnerability through different time of day (ii) No consideration of the characteristics of links (e.g., length, speed limit) (iii) No consideration of demand response to disruptive events
Almotahari and Yazici (2019)	Analytical/system-based	full network scan	(i) Low efficiency for large size networks (ii) No consideration of link interactions (iii) No realistic representation of the behavioural responses to disruption (static traffic assignment)
García-Palomares et al. (2018)	Analytical/system-based	full network scan	(i) Applied only to high capacity road network (e.g., freeways) (ii) No realistic representation of the behavioural responses to disruption (no traffic assignment)
Du et al. (2014)	Analytical/system-based	full network scan	(i) Proposed a simulation approach for logistics transport network (ii) No realistic representation of the behavioural responses to disruption (static traffic assignment: Dijkstra algorithm)
Knoop et al. (2012)	Analytical/system-based	pre-selected links	(i) Proposed a multi-linear fit of indicators as vulnerability index (ii) Relatively fast for large scale networks (iii) Not realistic consideration of incident occurrence probability (assumed proportional to flow)
Taylor and Susilawati (2012)	Analytical/system-based	pre-selected links	(i) Employed intelligent screening procedure to increase the computational efficiency (ii) More suitable for rural and remote areas (iii) Less suitable for congested urban network
Jenelius and Mattsson (2015)	Analytical/system-based	pre-selected links	(i) Assumed that the closure of a link does not affect the travel time on any other link (ii) Static and constant demand is considered during disruption (iii) Not suitable for congested urban areas
Chen et al., (2012)	Analytical/system-based	Impact area	(i) Introduces local vulnerability ranking (ii) More realistic representation of flow distribution after disruptive events (iii) Employs deterministic impact area approach with fixed number of affected links (iv) Not very realistic impact area determination
Erath et al. (2009)	Simulation/system-based	Impact area	(i) Modeled the impact of long-lasting events modeling direct and indirect impact of closure (ii) Considered only the impact of major links closure (iii) Assumed inelastic demand with respect to mode choice and destination choice shift

*Links are randomly selected based on random attacks to network elements

3.2 Network pace distribution

Previous vulnerability analysis measures the increase in the total network travel cost associated with each individual link closure to accordingly determine the global vulnerability ranking. In the analysis conducted in the literature, the impact of each link closure is measured for before and after closure by running a network traffic assignment without taking into account the duration of the closure and how quickly a link recovers from the incident or closure. In other words, it is not realistic to account for the impact of link closure by solely measuring the maximum possible increase in travel time. This single point in time measure (in terms of increase in the travel time only when the link is fully closed) does not fully consider the dynamic ramification of closure during the full-time span from link closure to recovery. This analysis is of interest as some of the links are more prone to incidents and simultaneously recover much more slowly from the incident compared to other links due to geometrical properties of the network in the vicinity of these links. Instances of geometrical and topological properties influencing such analysis include lower network redundancy in the vicinity of the links subject to incident; higher link betweenness; lack of alternative links to the links subject to incident; and one directional links subject to incident, among others. Multi-year observed travel time and incident (i.e., collision and link closure) data carry this information and can be mined to determine the full spatial and temporal impact of link disruptions in the studied network.

To identify the impact of an incident occurring in a given link j, and consistent with the finding of Erath et al., (2009) and Chen et al., (2012), it is assumed that the spatial and temporal impact of the closure of link j is limited to its impact area G', which is a subnetwork of the studied network G, and does not disperse throughout the whole network G. It is further

assumed that the immediate neighboring links (i.e., upstream and downstream neighboring links) are the most affected since the impact of an incident fades out as one moves farther from the link subject to incident. The experienced travel times during a disruptive event for each affected link in the impact area G' can be best represented as a distribution rather than an average travel time value. To determine the impact of a disruptive event on each link in the impact area G', it is not realistic to account for the increase in the travel time by directly subtracting the experienced travel time from the free flow speed. In fact, the expected travel time during different times of day can be considerably higher than the free flow travel time on a given link because of the within-day variation in the travel demand. Thus, to capture the impact of a disruptive event, the deviation between the mean of the historical distribution of pace (min/km) (i.e., normalized travel time defined as the ratio of link travel time to link distance) and the experienced pace during a disruptive event in each time period is calculated. Since the goal is to measure the impact of disruptive events on the performance of the subnetwork, it is more desirable to model the impact in terms of the change in pace as opposed to using the more traditional measure of vehicle count. Based on the shape of the fundamental diagram, when the vehicle flow drops significantly in a link, it is difficult to identify whether the underlying reason being simply low demand for travel or stop and go due to significant increase in congestion. The pace distribution clearly distinguishes between these possibilities since a smaller pace consistently represents decreased demand and higher pace indicates more congestion.

The pace of traffic on a link in a given network is equal to the inverse of Edie's generalized average speed (Edie, 1963). Therefore, the pace of a given link j in the set of links J in a hypothetical subnetwork G' in time period T and time interval τ where $\tau \in T$,

denoted by $P_T^\tau(j)$, can be calculated as Equation (3.1). The time interval τ is the granularity offered by the historical data. For instance, if the pace data is updated every 1 minute, $\tau = 1$ minute.

$$P_T^\tau(j) = \frac{u_T^\tau(j)}{l(j)} \tag{3.1}$$

where $u_T^\tau(j)$ is the travel time of link j in the time period T and time interval τ, and $l(j)$ is the length of link j. The traffic condition of the subnetwork G' consisting of $|J|$ links in the time period T and time interval τ can be then shown as the *mean pace vector*, denoted by $a_T^\tau(G')$ which consists of the pace of the links in the subnetwork G' as follows:

$$a_T^\tau(G') = \left(P_T^\tau(j)\right) \in \mathbb{R}^{1\times|J|} \tag{3.2}$$

The mean pace vector has a strong periodic pattern in the sense that not only the mean pace vector of the subnetwork G'has a significant time-of-day variations, but it also fluctuates from day to day. Such within-day and day-to-day dynamics in the mean pace vector can be modeled by grouping pace data into different clusters based on their time-of-day and time-of-week variation. Such grouping can be accommodated by the reference set $Q(G')$ which consists of all the reference sets constructed for different times of day across all available travel time data (i.e., $Q_T(G')$). Since there are 168 hours in week, the reference set $Q(G')$ is defined as below:

$$\begin{aligned} &Q(G') = \left(Q_{T,j}(G')\right)_{168\times|J|} \\ &Q_{T,j}(G') \in \mathbb{R}^{M_{T,j}\times 1} \end{aligned} \tag{3.3}$$

Where $M_{T,j}$ is the number of mean pace entries available for link j in time period T. The reference set $Q(G')$ is a matrix that its entries are themselves matrices $Q_{T,j}(G')$ that contains all the mean pace vectors in subnetwork G' which occur at the time period T. The grouping of the mean pace vectors based on their time-of-day and time-of-week is based on the assumption that the traffic conditions observed in a specific day of week and time period (e.g., Thursdays between 4-5 PM) on a specific link are fairly similar with minor deviation in the absence of disruptive events. When estimating how typical the traffic condition is in each of the 168 time periods in a week, the mean pace data during disruptive events (e.g., collisions, link closure, etc.) should not be included in the reference set $Q(G')$. The construction of the reference set $Q(G')$ makes it possible to calculate the expected mean pace vector of the subnetwork G', denoted by $\mu(G')$. $\mu(G')$ is a matrix consisting of the expected mean pace of subnetwork G' for each time period T (i.e., $\mu_T(G')$) and is defined as follows:

$$\mu(G') = \left(\mu_T(j)\right) \in \mathbb{R}^{168\times|J|}$$

$$\mu_T(j) = \frac{1}{M_{T,j}} \sum_{\tau=1}^{M_{T,j}} P_T^{\tau}(j) \tag{3.4}$$

The $\mu_T(G')$ is itself a matrix where its entries are the expected pace of each link j at the time period T, $\mu_T(j)$. In addition to the expected mean pace vector calculated by Equation (3.4), the PDF of the typical behaviour of all links in the network could be obtained for all the 168 possible time periods. Estimation of the typical behaviour of traffic for all links in the network make it possible to detect the deviation of the pace from the typical behaviour during unusual events.

Figure 3.1 shows a hypothetical subnetwork consisting of 10 directed links and 4 centroid nodes with 1 and 2 as origin and 8 and 9 as destination nodes. The link subject to incident (i.e., link connecting node 4 to node 6) is shown by a dotted line. For each link in the impact area, the typical behaviour in terms of the distribution of pace is shown as the blue curve and the red curve shows the pace distribution during a disruptive event. Once the distribution of the typical behaviour and distribution of the pace during a disruptive event is constructed for all links in the subnetwork, the deviation from the typical behaviour could be utilized to determine the spatiotemporal impact of a disruptive event. These procedures will be discussed in the subsequent sections.

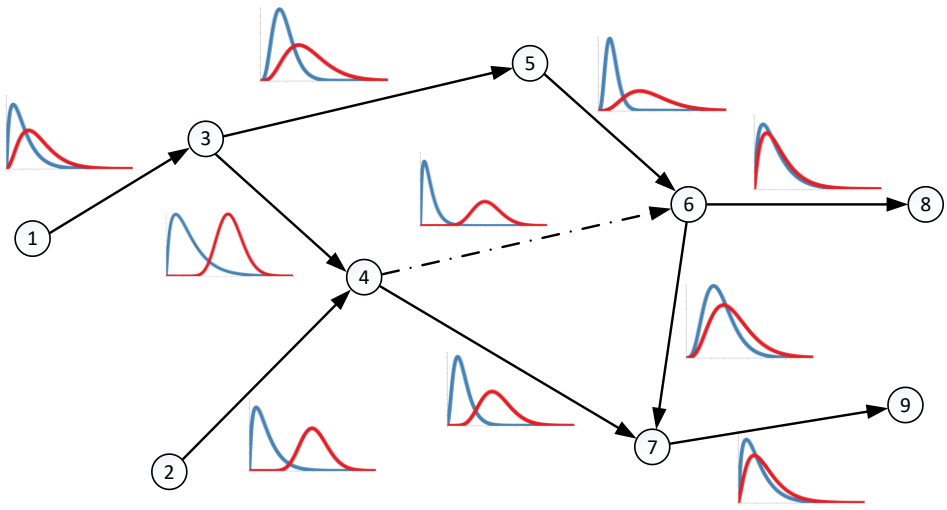

Figure. 3.1. Hypothetical impact area consisting of links affected by a disruptive event

3.3 Estimating the spatiotemporal impact of disruptive events

The mean pace vector of subnetwork G' is likely high in the presence of disruptive events. As the probability of disruptive events is very low, I am expecting that low mean pace occurs less frequently. Therefore, considering the distribution of the travel pace data

for a given link in the subnetwork G', very high mean pace data are located at the tail of the distribution. In the analysis in which the impact of the tail of the PDF is necessary to estimate, a classical central limit theorem may not be applicable to estimate the extent of the impact.

Various methods are employed in the literature to determine the spatiotemporal impact of disruptive events. Previous literature is mostly devoted to simulation and theoretical modeling approaches to determine the impact of disruptive events. More specifically, the literature can be divided into 3 groups depending on the deployed methodology: 1) theoretical modeling including deterministic queuing theory and kinematic shockwave theory (Asakura et al., 2017;), 2) field survey analysis (Sermons and Koppelman, 1996), 3) simulation studies (Li et al., 2020; Sethi et al., 1995). The outcomes of the previously mentioned approaches rely mainly on theoretical modeling of transportation networks or alternatively simulation-based analysis with none or little consideration of real-world traffic data. In addition, these events differ significantly in their duration, location and intensity; thus, their spatiotemporal impact events cannot be accurately estimated using these methods which restricts the spatial transferability of the developed models.

Recently, with emerging data collection technologies, high resolution traffic data is collected and mined and thus can be further explored to measure the spatiotemporal impact of disruptive events. This introduces the fourth method of estimating the incident impact which relies on empirical studies using real observed traffic data (Chakraborty et. al., 2019; Jeong et al., 2011; Pan et al., 2015; Sun et al., 2016; Liu et al., 2019). Different traffic parameters have been deployed by previous studies to estimate the spatiotemporal impact of collisions including traffic volume (Sun et al., 2016), travel speed (Chakraborty et. al., 2019;

Pan et al., 2015), combination of both traffic volume and travel speed (Liu et al., 2019), and combination of both travel speed and traffic occupancy (Jeong et al., 2011).

Relying on the assumption that the impact of an incident propagates over time and space, previous studies mainly focused on the impact of incidents on only the main corridor (i.e., the road segment directly affected by the incident) and its upstream sections. In reality, the effect of an incident does not impose restrictions on only the main corridor, but its impact disperses throughout the whole impact area consisting of the neighboring links (Chen et al., 2012). The sole consideration of the main corridor as the impact area leads to the underestimation of the spatial and temporal impact of incidents and thus, the obtained incident time might not reflect the correct time when traffic in the network is back to normal. In addition, because of the back-propagating shockwaves, the impact area follows a growing/shrinking pattern after the occurrence of an incident (Pan et al., 2015), therefore previous approaches are unable to capture the dynamic pattern of the impact area. Finally, the literature is mainly devoted to detecting incidents on either freeways (Chakraborty et. al., 2019; Jeong et al., 2011; Liu et al., 2019; Pan et al., 2015;) or arterials (Li et al., 2020; Sermons and Koppelman, 1996; Sethi et al., 1995) with little consideration given to estimating the spatiotemporal impact of incidents on other types of road simultaneously.

Recently, Donovan and Work (2017) studied the impact of 2012 Hurricane Sandy on the New York City road network. They constructed one mean vector including data related to both typical behaviour of traffic and during the hurricane for different types of trip. They estimated the temporal impact of the hurricane by estimating the Mahalanobis distance (Mahalanobis, 1936) across different time periods to measure the deviation of the pace during the hurricane from the typical pace of the network. Using Mahalanobis distance to

estimate the temporal impact of a disruptive event requires consideration of its appropriate threshold value above which the pace data are considered as extremes. The choice of threshold is tricky as the estimated temporal impact could be overly sensitive to the choice of threshold. Moreover, because of this sensitivity, this method might not be accurate enough to estimate the temporal impact of less-extreme events (e.g., collision, signal failure or any event leading to single link closure). The extent of impact could be determined more accurately by leveraging non-authoritative data and its integration with established data and methods. In this regard, Schnebele et al., (2014) evaluated the spatiotemporal impact of Hurricane Sandy by analysing non-authoritative data. Estimation of the temporal impact of these less-extreme events is necessary to identify the vulnerable elements (e.g., links and nodes) of a road network. In addition, modeling the impact of disruptive events could be inefficient if the pace PDF is constructed only once for all available mean pace vector data, especially in the case of light-tailed PDFs of the pace.

To overcome the above-mentioned limitations, an efficient algorithm was developed in this research to measure the spatiotemporal impact of disruptive events taking into account the impact of the event on all neighboring affected links. Tracking the evolution of the impact on neighbouring links enables us to capture the true temporal impact of these events. In addition, the dynamic behaviour of the spatial propagation of events is estimated in each time step considering the dynamic growing/shrinking pattern after occurrence of events. Moreover, the proposed model is a more general approach in the sense that the application of the proposed model is not restricted to only freeways and arterials. The developed model is also computationally efficient as its only input is the historical travel time data which reduces the computational burden of the proposed approach as opposed to the other methods

relying on various sources of data. Finally, the efficiency of the proposed model is enhanced by constructing two mean pace distributions; one for typical behaviour of the subnetwork G' and one for the mean pace during a disruptive event, as depicted in Figure 3.1, instead of finding the deviation of the pace from the typical behaviour by construction of only one pace PDF for the subnetwork G' as used in the literature (Donovan and Work, 2017).

3.3.1 Proposed spatiotemporal impact estimation model

Following the discussion in Section 3.2, two sets of pace distributions are constructed for the impact area G'; one set of mean pace distributions during a disruptive event and one set representing the typical mean pace distribution. In this section, the spatiotemporal impact of disruptive events is estimated using these two sets of mean pace distribution.

Let $e_k(g, \bigcup_{i=r}^{s} T_i),\ r \leq s\ \&\ r, s \in \{1, \ldots, 168\}$ denote the k^{th} single disruptive event associated with link g which is subject to incident. Since the impact of the disruptive event might not be limited to only one time period, the latter part (i.e., $\bigcup_{i=r}^{s} T_i$) indicates the time periods in which the impact of the disruptive event might persist. The set of disruptive events associated with the link g (i.e., $E(g)$) can be constructed as follows:

$$E(g) = \left\{ e_k \left(g, \bigcup_{i=r}^{s} T_i \right) \right\} \qquad k = 1, \ldots, |E(g)|\ \&\ r \leq s\ \&\ r, s \in \{1, \ldots, 168\} \tag{3.5}$$

The mean pace vector of the subnetwork G' during a disruptive event $e_k(g, \bigcup_{i=r}^{s} T_i)$ in time period T and time interval τ, (i.e., $a_T^{\tau}\left(G', e_k(g, \bigcup_{i=r}^{s} T_i)\right)$) is constructed as follows:

$$a_T^{\tau}\left(G', e_k\left(g, \bigcup_{i=r}^{s} T_i\right)\right) = \left(P_T^{\tau}\left(j, e_k\left(g, \bigcup_{i=r}^{s} T_i\right)\right)\right) \in \mathbb{R}^{|J|\times 1} \qquad r \le s \,\&\, r, s \in \{1, \ldots, 168\} \qquad (3.6)$$

where $P_T^{\tau}\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right)$ is the pace of link j during the event $e_k(g, \cup_{i=r}^{s} T_i)$ in time period T and time interval τ. The affected links in each time interval τ could be determined by comparing $a_T^{\tau}\left(G', e_k(g, \cup_{i=r}^{s} T_i)\right)$ with the typical pace behaviour of the subnetwork G'. More specifically, the following hypothetical test is performed at 95% confidence interval to construct the impact area G', which corresponds to those links for which the pace during the disruptive event is greater than the expected pace (i.e., $P_T^{\tau}\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right) \ge \mu_T(j)$).

$$\begin{aligned} &Null\ Hypothesis\text{: } H_0\text{: } P_T^{\tau}\left(j, e_k\left(g, \bigcup_{i=r}^{s} T_i\right)\right) = \mu_T(j) \\ &Alternative\ Hypothesis\text{: } H_1\text{: } P_T^{\tau}\left(j, e_k\left(g, \bigcup_{i=r}^{s} T_i\right)\right) \neq \mu_T(j) \end{aligned} \qquad r \le s \,\&\, r, s \in \{1, \ldots, 168\} \qquad (3.7)$$

In other words, links for which the pace during the disruptive event is significantly higher than the expected pace would be identified as affected by the event and thus is marked as one of the links in the impact area G'.

The hypothesis tested in Equation (3.7) is repeated for every time interval in each of the time periods, T_i, associated with the disruptive events to determine the dynamics of the growing/shrinking patterns of the impact area G' during the disruptive event. The impact area G' is then constructed for all available time intervals of τ. Once the impact area is constructed for every time interval τ, the duration of the disruptive event $e_k(g, \cup_{i=r}^{s} T_i)$ (i.e., $U\left(e_k(g, \cup_{i=r}^{s} T_i)\right)$) could be obtained relying on the assumption that the mean pace vector

of the impact area G' for different time intervals remains fairly similar. In other words, the temporal impact is assessed assuming that the traffic congestion patterns through different time intervals during the incident have greater similarities between each other compared to the traffic pattern before the time the event occurs and after the time the traffic condition is back to normal.

In this study, by the temporal impact of an incident I mean the difference between the incident occurrence time (T_0) and flow *recovery time* (T_7) which is shown as *time to return to normal flow* in Figure 3.2. The flow recovery time is when the impact of an incident is totally dissipated and the flow in the impact area returns totally to normal condition. In traffic incident reports, the reported incident time of an incident (T_1) is usually recorded instead of the incident occurrence time. In addition, either roadway clearance (T_5) or incident clearance time (T_6) are usually reported as the end time of an incident. Thus, the difference between the reported start and end time does not fully represent the temporal impact of a given incident as even if the incident is removed its impact can still be propagating on other links by space and by time.

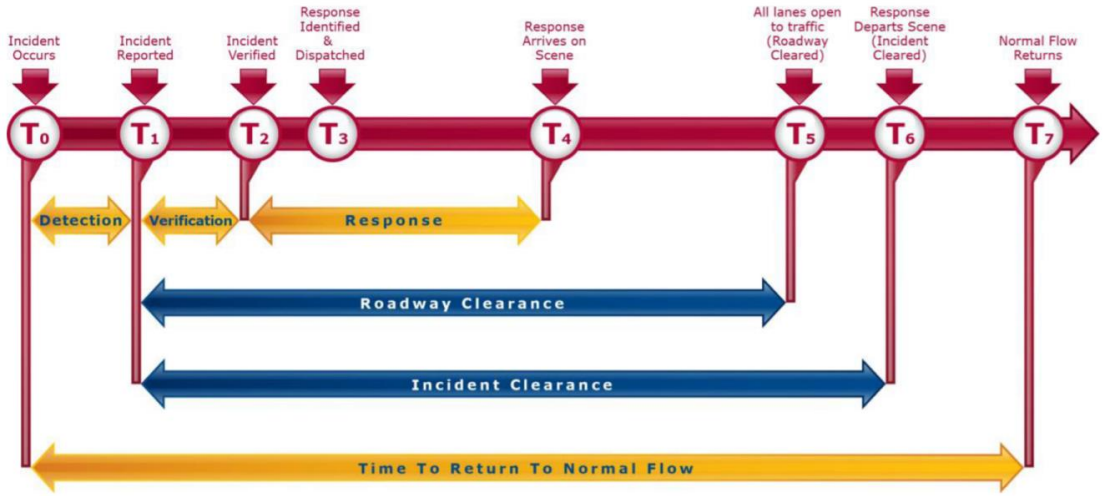

Figure 3.2. Timeline of a traffic incident (Amer et al., 2015)

To make the comparison between the mean pace vectors of the impact area G' during the disruptive event easier, all pace data are labelled and clustered. The primary data for the clustering analysis is the directed graph of the links which are connected in space with their immediate neighbouring links (i.e., the upstream and downstream links connected to the affected link) and in time with the pace data associated with the previous and next time intervals for the affected link. Each observation in the impact area G' is shown by start and end node of the corresponding link j, (i.e., $(x_{j,s}\ y_{j,s})$ and $(x_{j,e}\ y_{j,e})$ respectively), time interval τ and its pace during time interval τ (i.e., $P_T^{\tau}\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right)$).

The entire database consisting of links and their corresponding pace distribution during different time intervals throughout the whole incident period is synthesized into individual observations (i.e., a couple composed by a link and its pace deviation from the typical behavior during a time interval shown as $\left(j, P_T^{\tau}\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right) - \mu_T(j)\right)$). Since basic information that is reported on incidents such as their occurrence time and clearance time usually comes from the perceptions of those involved in the incident (also called non-authoritative data) or a guess from the officer reporting the incident, information on the incident occurrence and clearance times might not be accurate enough to derive the temporal impact of incidents. Therefore, the boundary of $B\left(e_k(g, \cup_{i=r}^{s} T_i)\right) = \left[ST_{reported} - 30\ min, ET_{reported} + 30\ min\right]$ is chosen for the aim of clustering, where $B\left(e_k(g, \cup_{i=r}^{s} T_i)\right)$ is the time boundary corresponding to the incident $e_k(g, \cup_{i=r}^{s} T_i)$ and $ST_{reported}$ and $ET_{reported}$ are its reported occurrence time and clearance time respectively. These intervals are set to obtain the true temporal impact of the incident from the real occurrence time of the incident to the true time the impact of the incident is cleared from the road network.

Once the set of all observations within the boundary $B\left(e_k(g, \cup_{i=r}^{s} T_i)\right)$ is constructed, Ck-means clustering method is performed to assign a cluster ID to each of the observations in the set. Ck-means is the improved version of the k-means clustering method originally developed by Wang and Song (2011). It has been shown recently that k-means is an efficient clustering method in partitioning urban road networks considering both spatial locations of links and their corresponding travel time information (Ji and Geroliminis, 2012; Lopez et al., 2017). However, the disadvantage of k-means algorithm is that it depends heavily on the initial cluster centers (i.e., number of clusters) which is required as the clustering input by the user and thus not always optimal nor repeatable. Moreover, the runtime of the k-mean algorithm increases exponentially with the number of clusters. Ck-means algorithm overcomes these issues by using a dynamic programming algorithm to guarantee optimality of clustering by minimizing the sum of the squares of within-cluster distances from each observation to its corresponding cluster mean. As the result, the runtime of the Ck-means algorithm is a linear function of the number of clusters which makes it more efficient than K-means when the number of clusters is high. The reader is referred to Wang and Song (2011) for more information on the Ck-means algorithm.

After performing the Ck-means algorithm on all observations $\left(j, P_T^{\tau}\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right) - \mu_T(j)\right)$ within the boundary $B\left(e_k(g, \cup_{i=r}^{s} T_i)\right)$, each observation is assigned a cluster ID. Each set of clustered observations belonging to the same time interval τ is then synthesized into a single ordered vector of all observations π_t, whose values are the cluster IDs. To compare two consecutive vectors of observations which belong to two consecutive time intervals τ and τ' (i.e., π_τ and $\pi_{\tau'}$) and to assess how similar are these

vectors, I use normalized mutual information (NMI) which measures the proximity between two vectors obtained from clustering (Strehl and Ghosh, 2002) as follows:

$$NMI(\pi_\tau, \pi_{\tau'}) = \frac{I(\pi_\tau, \pi_{\tau'})}{\sqrt{H(\pi_\tau)H(\pi_{\tau'})}} = \frac{H(\pi_\tau) + H(\pi_{\tau'}) - H(\pi_\tau, \pi_{\tau'})}{\sqrt{H(\pi_\tau)H(\pi_{\tau'})}} \quad (3.8)$$

Where $NMI(\pi_\tau, \pi_{\tau'})$ is normalized mutual information between observation vector π_τ and $\pi_{\tau'}$, $I(\pi_\tau, \pi_{\tau'})$ denotes the mutual information between π_τ and $\pi_{\tau'}$ which measures the mutual dependence between two random variables. $H(\pi_\tau)$ and $H(\pi_{\tau'})$ are the entropy of π_τ and $\pi_{\tau'}$ respectively and $H(\pi_\tau, \pi_{\tau'})$ is the joint entropy of π_τ and $\pi_{\tau'}$. The similarity between every two consecutive vectors of observations corresponding to two consecutive time intervals within the boundary $B\left(e_k(g, \cup_{i=r}^{s} T_i)\right)$ is then calculated and plotted to spot any unusual variations in the value of NMI which indicates dissimilarity between the observation vectors of the two consecutive time intervals from which the occurrence and flow recovery time of the event could be inferred. It is to be expected that NMI values for time intervals before the occurrence of the incident remain fairly high followed by a sharp drop in the time of incident occurrence. NMI values are also expected to remain high throughout the duration of the incident followed by a drop when the incident is cleared from the road network and its impacts start to dampen.

3.4 Impact area vulnerability analysis approach

As discussed previously in Section 3.1.2.2, the common approach to determine the global vulnerability of links in a given road network is the full scan network approach. In this approach, links in a given road network are removed successively and network traffic

assignment is carried out to evaluate the travellers' rerouting response to the link closure and consequently measure the impact of link removal on the total network cost. The global vulnerability ranking is then determined by sorting the value of the increase in the total cost of the network in descending order. This approach is not without limitations as it is computationally intensive for large-scale networks. Second, empirical findings show that the impact of link closure imposes restrictions only on an impact area in the vicinity of the link subject to incident; thus, the impact does not disperse throughout the whole network (Danczyk and Liu, 2010). In a few studies where the impact area approach is carried out to determine the impact area vulnerability ranking, the size of the impact area, in terms of the level of proximity of the affected links to the link subject to incident, was given as the model input rather than the output of the model (Chen et al., 2012). In other words, in previous studies, it was assumed that the size of the impact area is fixed and independent of the type of the link subject to incident and the network characteristics in the vicinity of that link. The limitations of previous analytical approaches necessitate conducting more empirical analysis to evaluate how the impact area resulting from a link failure propagates through time and how the characteristic of the link subject to incident contributes to the magnitude of the impact. Finally, the focus of previous approaches was mainly on the consequences of the disruptive events to determine the vulnerable elements in a given road network with no consideration of the probability of disruptive events. In this section, a data driven impact vulnerability analysis approach is proposed using the empirical information regarding the spatiotemporal impact of the historical disruptive events to determine the impact area vulnerability ranking of links in a given road network considering both the probability and consequences of disruptive events.

3.4.1 Modeling probability and consequences of disruptive events

Once the spatiotemporal impact of a disruptive event is determined, the impact of the event could be assessed by measuring the deviation of the mean observed pace of a link during the incident from the typical pace of that link and also by evaluating how likely the pace during the disruptive event is higher than the typical pace of that link in the impact area G'. The former measure represents the consequence of failure on a given link – whether it is a link subject to incident or an affected link- and the latter measure denotes the probability that a given link is significantly affected by an incident. The probability that the pace during an incident on a given link is greater than the typical pace of that link can be estimated by applying the *limit state function* (Melchers and Beck, 1999), which is usually applied in structural reliability analysis.

The probability that the pace of link j during an incident $e_k(g, \cup_{i=r}^{s} T_i)$ (i.e., p_f) is greater than the nominal pace of link j could be shown as below:

$$p_f = P\left(P_{\cup_{i=r}^{s} T_i}(j, N) \leq P_{U\left(e_k(g, \cup_{i=r}^{s} T_i)\right)}\left(j, e_k\left(g, \bigcup_{i=r}^{s} T_i\right)\right)\right) \qquad \begin{matrix} k = 1, \dots, |E(g)| \;\&\; r \\ \leq s \;\&\; r, s \in \{1, \dots, 168\} \end{matrix} \tag{3.9}$$

where $P_{\cup_{i=r}^{s} T_i}(j, N)$ is the nominal pace of link j during the time periods $\cup_{i=r}^{s} T_i$ associated with the incident $e_k(g, \cup_{i=r}^{s} T_i)$ and $P_{U\left(e_k(g, \cup_{i=r}^{s} T_i)\right)}\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right)$ shows the pace of link j during the temporal impact of the incident.[6] Denoting the marginal PDF of pace of link

[6] From now on, for the aim of simplicity, $P_{\cup_{i=r}^{s} T_i}(j, N)$ and $P_{U\left(e_k(g, \cup_{i=r}^{s} T_i)\right)}\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right)$ will be represented by $P_N(j)$ and $P_Q(j)$ respectively.

j during the temporal impact of the incident as $f_Q(j,q)$, and the marginal PDF of the nominal pace (i.e., typical pace) of link j during the time periods $\bigcup_{i=r}^{s} T_i$ associated with the incident as $f_N(j,n)$, Equation (3.9) can be re-written as follows:

$$p_f = P\big(P_N(j) - P_Q(j) \le 0\big) = P\big(G_{N,Q}(j) < 0\big) = \iint_D f_{N,Q}(j,n,q)\, dn\, dq \qquad (3.10)$$

where $f_{N,Q}(j,n,q)$ shows the joint (bivariate) density function corresponding to link j and D denotes the domain of the pace in which the pace during the incident is higher than the nominal pace which is represented by the hatched domain D in Figure 3.3. $G_{N,Q}(j)$ is termed the limit state function and the probability that link j is affected by the event is identical with the probability of limit state violation. In Figure 3.3, $\bar{P}_N(j)$ and $\bar{P}_Q(j)$ are the mean of $f_N(j,n)$ and $f_Q(j,q)$ respectively.

In essence, finding the probability that link j is affected by the incident on link g requires solving a multi-dimensional integration. However, finding a closed-form expression for the integration in Equation (3.10) might not be always possible given that the marginal PDFs $f_N(j,n)$ and $f_Q(j,q)$ might follow various statistical distributions. For the aim of finding this probability, simulation approaches (e.g., direct sampling methods) could be performed. In direct sampling methods, random samples of vectors of the random variables (i.e., nominal pace and during incident pace) are generated and the samples of the limit state function $G_{N,Q}(j)$ are obtained consequently.

In this research, to find the marginal PDFs of the nominal pace and the pace during incident, various PDFs are fitted to both the nominal pace data and incident pace data of link j. The choice of distributions is limited to the parametric distributions that are found by the

literature to properly describe the observed travel time data; normal (Zhong et al., 2020), lognormal (Herman and Lam, 1974; Richardson and Taylor, 1978; Pu, 2011), betta (Polus, 1979; Castillo et al., 2012), gamma (Herman and Lam, 1974; Polus, 1979; Kim and Mahmassani, 2015), exponential (Talley and Becker, 1987; Noland and Small, 1995) and generalized extreme value distribution (Al-Deek and Emam, 2006). Both MLE method and Bayesian inference are utilized in the literature to obtain the best fitted distributions of the empirical nominal pace and incident pace data. In the MLE method, the best fitted distribution is chosen by measuring the goodness of fit of all the above-mentioned distributions and comparing their maximum likelihood values. Once the best distributions according to both methods of parameter inference methods (i.e., MLE and Bayesian) are determined, various statistical techniques are employed to determine the superior distribution. More specifically, the fit of the CDF of the empirical pace data are compared to the CDFs obtained from parameter inference methods. Additionally, the quantile-quantile (q-q) plots and also the mean of the modeled and the mean of the empirical pace data are compared to find the best fitted distribution of the empirical nominal pace and incident pace data.

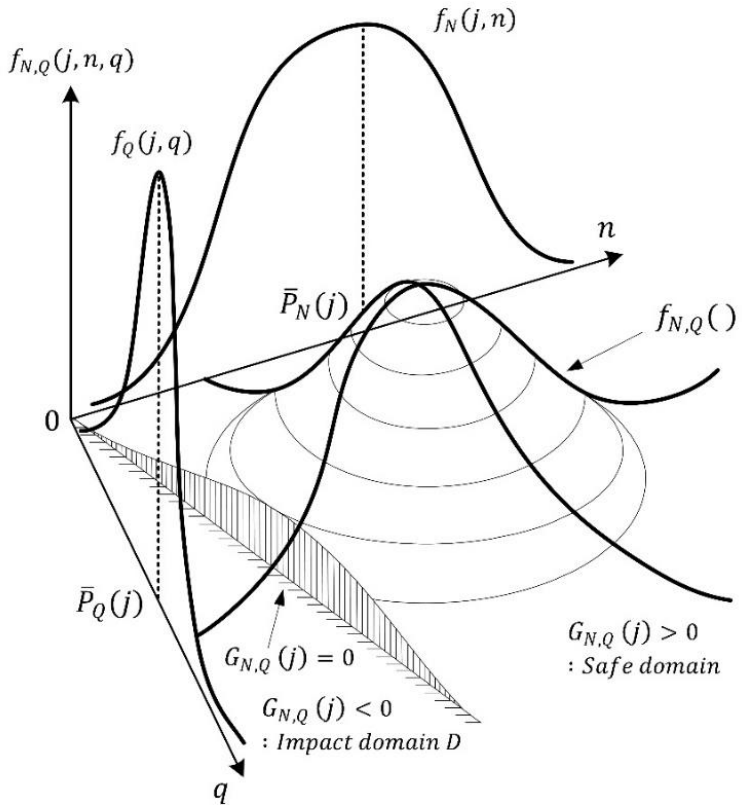

Figure 3.3. Impact domain D, marginal PDFs and the joint PDF in 3D

Depending on the underlying fitted distribution a closed-from solution or appropriate analytical approximation might not always exist. Therefore, in the next step, a crude Monte Carlo simulation approach is considered to measure the probability that a given link j is significantly affected by the incident $e_k(g, \cup_{i=r}^{s} T_i)$. To this end, a large equal number of nominal pace and incident pace samples (i.e., n_s) are generated according to each underlying distribution. For each pair of generated samples, the sample pace during the incident is subtracted from the nominal pace to obtain the value of the limit state function $G_{N,Q}(j)$ for each sample pair. The probability that link j is affected by the incident is then calculated by measuring the proportion of the samples in the domain D to the total number of experiments:

$$\phi_f(j) \approx \frac{n\left[G_{N,Q}\,(j) < 0\right]}{n_s} \tag{3.11}$$

where $n\left[G_{N,Q}\,(j) < 0\right]$ represents the number samples in the domain D. The required number of samples generated for each distribution depends on the underlying distributions and the desired accuracy of the estimated probability $\phi_f(j)$. To have reasonably high number of samples in the domain D, 10^5 samples for each distribution are generated initially. The number of samples in the domain D is then counted. The experiment will be performed again with higher number of experiments (i.e., 10^6) in the case that the number of samples in the domain D is not large enough.

3.4.2 Impact area vulnerability index

Previous studies mainly focused on finding the vulnerable links by only measuring the consequence of failure of given link j with no consideration given to its possible impact resulting from the failure of adjacent links. In fact, due to the possibility of the existing correlation between adjacent links (Chen et al., 2011; Chen et al., 2012), failure of a given link j could lead to the failure of other links in its vicinity or at least it can affect the travel pace of these links. Of course, the characteristics of the link itself in terms of capacity and the time of the occurrence of the disruptive event play a major factor in whether the impacts are propagated to other links or not. For instance, a collision blocking two lanes in a major freeway might result in a different size of the impact area depending on when the collision occurs. While if the collision occurs during the night, the resulting traffic disruption can be easily absorbed, during the morning or afternoon peak hour, the diverted traffic and possibly

blocked on/off ramps would significantly affect other neighboring roads/arterials; possibly inducing a propagating spillback domino effect on other links.

To take into consideration both the impact of incident occurring on link j on the performance of both link j itself and other links in its vicinity and also the impact of incidents occurring on other links in the vicinity of link j on the performance of link j, the probability vector of link j (i.e., $\phi(j)$) is calculated using the procedure in Section 3.4.1. This probability vector itself is constructed by two smaller probability vectors; $\phi\left(j, e_k(j, \cup_{i=r}^{s} T_i)\right)$ contains the probabilities that link j is affected by incidents on link j itself (direct impact probabilities) and $\phi\left(j, e_k(g, \cup_{i=r}^{s} T_i)\right)$ which includes the probabilities that link j is affected by incidents occurring on other adjacent links (indirect impact probabilities). A new vulnerability index (VUL_j) is proposed considering both the direct and indirect impact probabilities on link j and the consequence of the impact in terms of increase in the mean pace of link j.

$$VUL_j = a(j)\left(\sum_{k=1}^{m_j}\sum_{h=1}^{n_G}\alpha_{h,j}\phi\left(h, e_k\left(j, \bigcup_{i=r}^{s} T_i\right)\right)\left(\bar{P}_Q(h) - \bar{P}_N(h)\right) + \sum_{g=1}^{n_G}\sum_{k=1}^{m_g}\alpha_{j,g}\phi\left(j, e_k\left(g, \bigcup_{i=r}^{s} T_i\right)\right)\left(\bar{P}_Q(j) - \bar{P}_N(j)\right)\right) \quad (3.12)$$

where m_j and m_g are the total number of disruptive events that occurred on link j and link g respectively, and n_G shows the total number of links in the network G. $\alpha_{h,j}$ and $\alpha_{j,g}$ are binary parameters defined as below:

$$\alpha_{h,j} = \begin{cases} 1 & \text{If link } h \text{ is affected by incident on link } j \\ 0 & \text{If link } h \text{ is not affected by incident on link } j \end{cases} \quad (3.13)$$

$$\alpha_{j,g} = \begin{cases} 1 & \text{If link } j \text{ is affected by incident on link } g \\ 0 & \text{If link } j \text{ is not affected by incident on link } g \end{cases} \quad (3.14)$$

In Equation (3.12), $a(j)$ indicates the importance of link j in network G. Adopting the idea of trip importance introduced by Zhong et al. (2019), a new link importance measure (i.e., $a(j)$) is introduced in this study. Given n_G links, denote $P_T(j,i) = \frac{u_T(j,i)}{l(j)}$ as the pace distribution of link $j = 1, \dots, n_G$ in time period $T = 1, \dots 24$, of day i and $R_T(G,i) = \frac{\sum_{j=1}^{n_G} l(j) P_T(j,i)}{\sum_{j=1}^{n_G} l(j)}$ as the network-scale pace distribution of network *G* at time period *T* of day i, where $u_T(j,i)$ is the travel time of link j in the time period *T* of day i.

A similar optimization problem to the one in Chapter 2 equation 2.32 is formulated to identify the importance of each link on network performance and is repeated for the reader's convenience as follows:

$$Z = min \sum_{i=1}^{N} \sum_{T=1}^{24} \left(E\left[R_T(G,i) - \sum_{j=1}^{n_G} a(j).P_T(j,i) \right] \right)^2 \quad (3.15)$$

$$s.t. \quad \sum_{j=1}^{n_G} a(j) = 1$$

$$0 \le a(j) \le 1$$

The results obtained from solving the above equation is a real number $a(j)$ between 0 and 1 for each link in the network. In other words, Equation (3.15) assigns an importance index $a(j)$ to each link by minimizing the deviation between the network-wide measure of pace distribution and the summation of the product of $a(j)$ with the individual pace distribution of links in network G. Since the objective function in Equation (3.15) is always positive, the optimization problem tries to assign higher $a(j)$ values to links with higher $P_T(j,i)$. In other words, the outcome of optimization problem in Equation (3.15) is assigning higher values of $a(j)$ to links with higher travel paces (e.g., major roads including freeways and arterials). Subsequently links with lower travel pace (e.g., minor links or links with relatively high level of service and operation conditions) are penalized by lower value of $a(j)$. An increase in $a(j)$ portrays a more significant contribution of link *j* to the overall network performance; thus, links with the highest values of $a(j)$ have the most substantial contribution to network performance.

The optimization problem in Equation (3.15) is a convex quadratic optimization problem (QP) with linear constraints which can be solved using exact optimization methods. The optimization problem is formulated and solved in Python using Gurobi solver (Gurobi Optimization, 2020). Using exact methods to solve the optimization problem in Equation (3.15) ensures reaching the global optimum with reasonable computation time.

3.5 Analysis of the incident and travel time data: Calgary network case study

3.5.1 Data description

Data collected from a real-world case study is analyzed to showcase the performance of the proposed spatiotemporal impact estimation model and its application in estimating the

link vulnerability. Travel time and incident data collected from various roadways in the City of Calgary over multiple years were analyzed. Around 85% of the Calgary road network, consisting of 5104 links, are selected for the aim of the analysis. The performance of the proposed approach is thus evaluated for the majority of links in the Calgary road network.

For the aim of this study, the travel time data of the selected corridors were collected for each single day between the end of December 2013 to the end of January 2019 and aggregated over 1-minute time intervals to fully investigate the impact of disruptive events on the performance of the links. The time interval of a minimum of 1 minute is required in order to have a sufficient number of travel times among which the impact of disruptive events is determined. The data was collected by INRIX Roadway Analytics which provides services that allow the users to view and query regional traffic flow information. The coverage of the INRIX Roadway Analytics includes provincial highways, the Major Road Network and select arterial roads within the Metro Calgary boundary.

Another dataset used in this study is the incident database which includes collision data from the City of Calgary for the same period of time as the travel time database (i.e., end of December 2013 to the end of January 2019). Information reported for each incident includes, the type of incident (i.e., collision, road repair, signal failure), date, time of the incident, location of the incident in the form of the latitude and longitude, incident severity in terms of the number of blocked lanes and reported occurrence time and clearance time of the incident. It is to be noted that road repair events were not considered in the analysis as they fall under planned events, which are out of scope of this research. The occurrence and clearance time of the incidents are mostly reported by either those involved in the incident or police officers. For a small portion of the historical incidents the occurrence time is

identified using traffic cameras in the vicinity of the incidents, however, the source of the reported occurrence time of the incident is not recorded in the database.

3.5.2 Investigating the spatiotemporal impact of incidents

The City of Calgary's incident database used in this study includes 30,814 individual incidents. Since the duration of work zones and other road works can last for months, they have been removed from the analysis performed to determine the spatiotemporal impact of incidents. After filtering those events, 12,504 events were included and the analysis was performed regardless of their severity (i.e., number of lane blockages). The incident database is integrated with the travel time database and incidents are assigned to their corresponding links by matching the coordinated in the both databases. Collisions occurring at intersections are assigned to their immediate neighbours (i.e. immediate upstream and downstream links) to have a large enough link for analysing the spatiotemporal impact of incident at intersections. The spatiotemporal impact analysis was implemented using Python 3.7.5 software on a personal computer with 3.4 GHz Intel Core^{TM} i7-6700 processor and 16GB of memory. The spatiotemporal impact determination took in average 30 second for each collision. To show the performance of the proposed approach in evaluating the temporal impact of the incident occurrences, results from 6 incidents from various roads in the City of Calgary are illustrated here. The characteristics of the incidents are shown in Table 3.2.

Table 3.2. Incident information including reported occurrence time, clearance time, estimated occurrence time and flow recovery time

Incident	Link subject to incident	Direction	Length (km)	Reported occurrence time	Reported clearance time	Estimated occurrence time	Estimated flow recovery time
1	Macleod Tr N (btw 6 Ave SE and 5 Ave SE)	NB	0.11	9:05	10:40	8:59	10:47
2	Crowchild Tr NW (btw Kensington Rd NW and Memorial Dr NW)	SB	0.21	8:58	11:20	8:59	11:24
3	Glenmore Tr SE (btw Deerfoot Tr SE and Blackfoot Tr SE)	WB	0.40	9:54	11:50	9:56	12:03
4	Canyon Meadows Dr SW (btw Intersection with Macleod Tr N)	WB	0.21	13:05	14:45	12:59	14:53
5	Macleod Tr N (btw Intersection with Midlake Blvd SE)	NB	0.23	17:10	17:40	17:02	17:46
6	Crowchild Tr NW (btw 53 St NW and Sarcee Tr NW)	WB	0.81	8:55	10:10	8:46	10:16

The temporal impact information reported in Table 3.2 is obtained by plotting NMI versus the 1-minute time intervals for incidents listed in Table 3.2 as depicted in Figure 3.6.

The occurrence time and the flow recovery time of the incidents are then estimated by spotting any sudden drop in the value of NMI which indicates dissimilarity between the observation vector of the two consecutive time intervals. The inference of the incidents' occurrence and flow recovery time is based on the assumption that NMI values remain fairly high for a short period of time, where the demand variation is negligible before the occurrence of the incident, followed by a sharp drop during the time of incident occurrence. NMI values are also expected to remain high for the incident impact duration followed by

another drop during the vicinity of the time that the impact of the incident occurring on the link is cleared from the impact area. The drops in the NMI values are indicated by red and green circles in Figure 3.6, to identify the inferred occurrence and the flow recovery time associated with each incident shown in Table 3.2.

To elaborate further on how the spatiotemporal impact of incidents are determined, the spatial impact of Incident 1 is determined by applying the approach developed in Section 3.3.1 and is represented in an example in Figures 3.4 and 3.5. Important time points including reported start time, estimated start time, reported clearance time and estimated flow recovery time are also shown in Figure 3.5. The link subject to incident is indicated by the black star and the affected links in each time interval are shown by dashed lines. In addition, the links that belong to the same cluster obtained from conducting Ck-means clustering technique discussed in Section 3.3.1, are shown by the same colour. In this specific incident, the outcome of the Ck-means clustering is 4 clusters that are shown in 4 different colours; blue, green, yellow and red. The mean pace of links in different clusters has an ascending order from the colour blue to red, while links within a same cluster (shown by same colour in Figure 3.5) have relatively similar paces From Figure 3.5, therefore it is evident that the impact of the incident has been dispersed throughout the specified part of Calgary downtown area depicted in Figure 3.4 well ahead of the reported start time. It is also shown that even after the clearance of the incident from the link subject to incident, the impact of the incident remains on some of the links in terms of the deviation of the pace from nominal pace of those links.

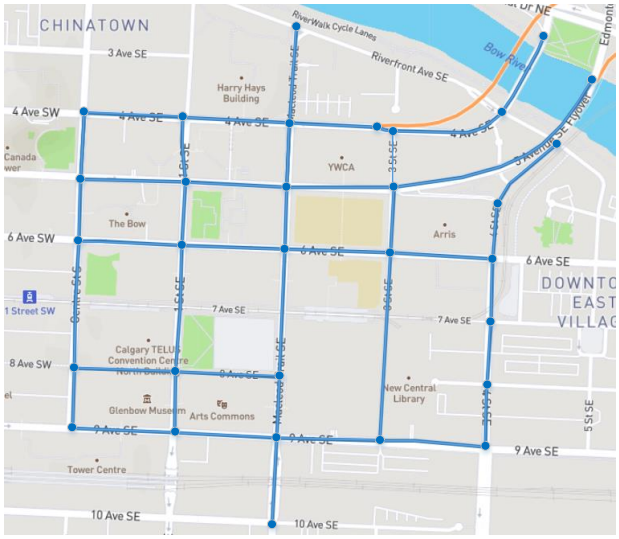

Figure 3.4. Part of Calgary downtown area affected by Incident 1

Based on the information in Table 3.2, the model estimated occurrence times are usually detected ahead of the reported occurrence time, while the estimated flow recovery time is always after the reported clearance time. Possible sources of these deviations could be error or delay in recording the incident occurrence and clearance time, which is evident in most cases in Table 3.2 where the occurrence and clearance times are rounded to the nearest 5 minutes.

To investigate whether the same trend is applied to other historical incidents, the spatiotemporal impact of all historical incidents was estimated and the deviation of the estimated quantities (i.e., occurrence time, flow recovery time and impact duration of the incident), from the reported ones is shown in Figure 3.7. The boxplot represented in Figure 3.7 includes minimum, maximum, median, mean (shown by plus sign), first quartile, and third quartile for the difference between the estimated and reported occurrence time, reported clearance time and flow recovery time, and time difference between reported incident

duration and estimated incident impact duration. Individual data is also plotted for each category. The mean of the difference between estimated and reported occurrence time is close to zero and in most cases the difference is negative, which indicates that the occurrence time is detected earlier by the proposed approach compared to the reported occurrence time. This finding is reasonable as the majority of the reported occurrence times are not accurate since they are usually reported by those involved in the incidents or by operators and sometime after the real occurrence time of the incident. The difference in the occurrence time could be positive though for some events, as indicated in Figure 3.7. The positive difference is justified by the fact that the developed methodology in Section 3.3 determine the temporal impact by comparing the pace pattern of the impact area. It might take the algorithm some time to recognize the change in the pace pattern throughout the impact area as the change in the pace after the incident could be gradual.

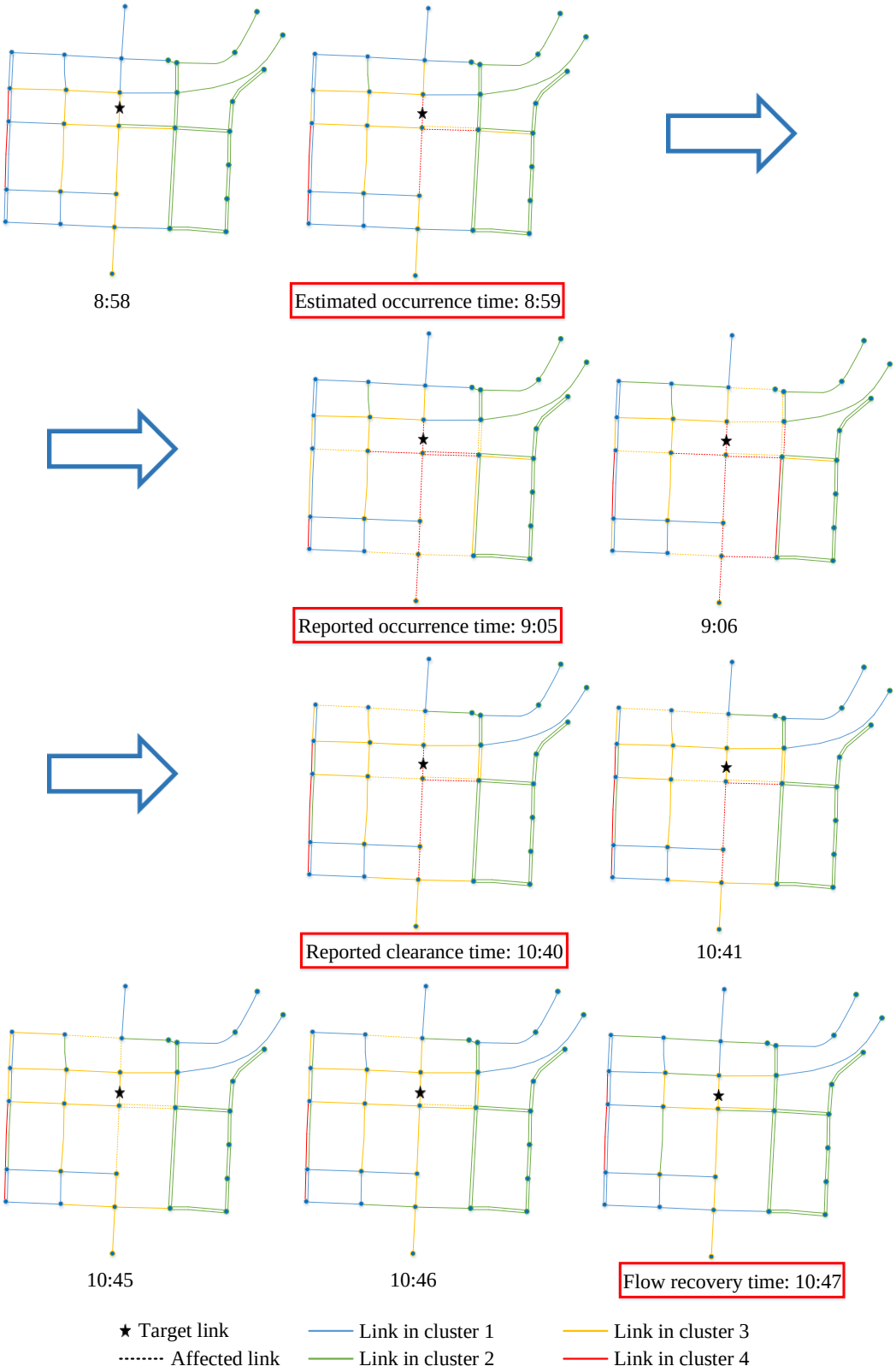

Figure 3.5. Spatial impact of Incident 1 in different time intervals

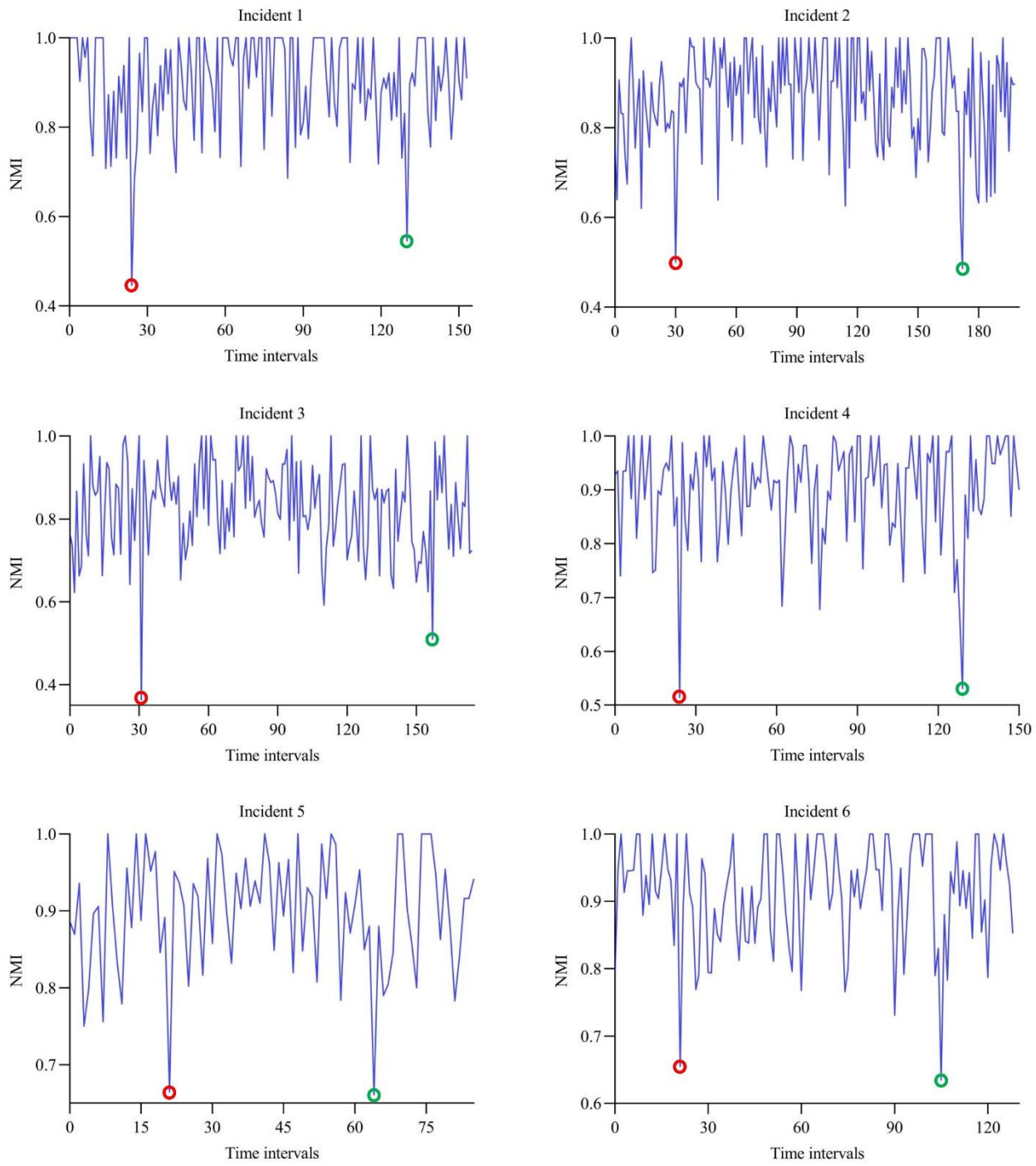

Figure 3.6. NMI values versus time intervals for selected collisions in Calgary road network

In contrast to the occurrence time, the majority of the differences between the flow recovery time and clearance time of incidents are positive, which shows that the real impact of incidents endures much longer in the network. This finding is intuitive since the recorded clearance time in the dataset is the time that a given incident is removed from the link subject to incident and is not reflective of the ongoing propagation of the shockwave induced by the incident on the links in the vicinity of the link subject to incident. Because of the relatively high difference between the reported clearance and estimated flow recovery time of the incidents, the estimated duration of the effect of incidents is found to be higher than the reported duration of the incident. In addition to the higher mean difference between the reported clearance time and estimated flow recovery time compared to the mean difference between the reported and estimated occurrence time (7.31 min versus 2.44 min), higher variation is observed in the difference of the reported clearance time and estimated flow recovery time. The $SD = 5.98$ min for the difference between the reported clearance time and estimated flow recovery time – compared to the $SD = 4.08$ min of the difference of the reported and estimated occurrence time, implies that the reported clearance time is not representative enough of the actual incident flow recovery time. While it is possible to record the occurrence time of the incident accurately by field observation, the flow recovery time of the incident from the impact area cannot be directly detected by field observation as the impact of the incident pertains even after the incident is removed from the link subject to the incident. The proposed approach can determine the enduring spatiotemporal impact of incidents which is crucial in road safety, resilience and road vulnerability analysis.

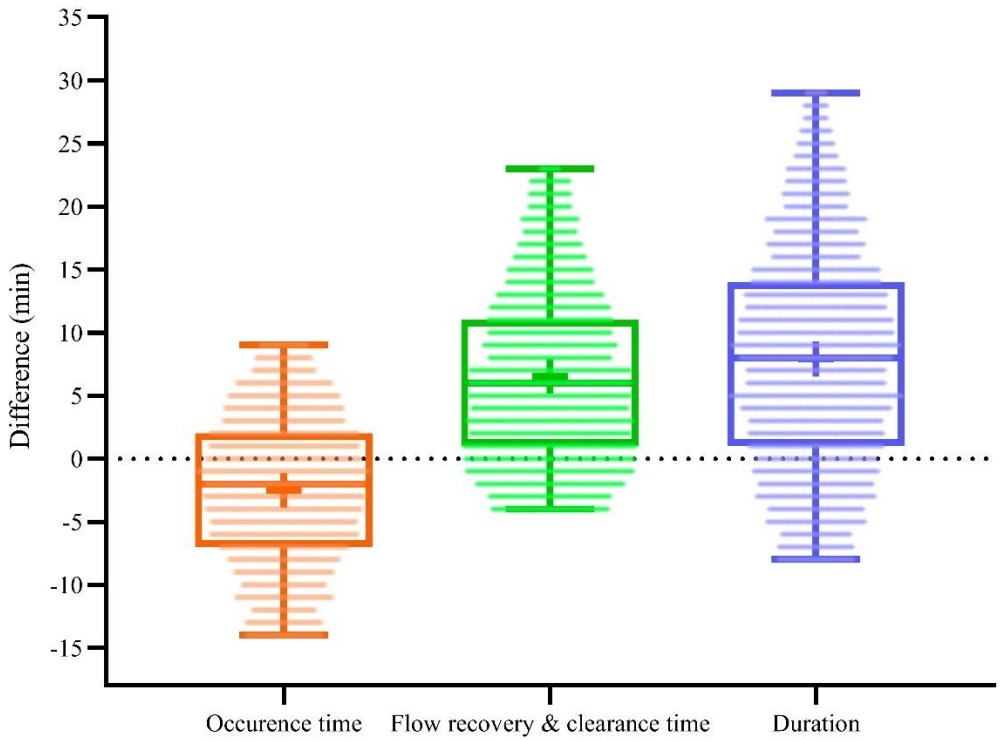

Figure 3.7. Time difference between the estimated and reported occurrence time, reported clearance time and flow recovery time and time difference between reported incident duration and estimated incident impact duration

It is possible to validate the proposed approach by comparing the estimated and the reported occurrence time when accurate reported occurrence time is available. Trustworthy source of this information could be the occurrence times extracted from CCTV traffic cameras. However, since the source of the records is not specified in the examined dataset, it is not possible to validate the model. The validation could be performed upon availability of such data by comparing the estimated and reported occurrence times of such incidents.

3.5.3 Calgary road network vulnerability analysis

Once the spatiotemporal impact of all incidents in the incident database of the City of Calgary is evaluated, the individual vulnerability index of each link is calculated using Equation (3.12). The vulnerability index accounts for both the direct and indirect impact of incidents affecting a given link *j*. By direct impact, I mean the impact of all incidents in which link *j* is the link where the incident occurs and the impact of link *j* failure is measured on the associated impact area. By indirect impact, I mean the impact of all incidents in which link *j* is not the link where the incident occurs but is affected by incidents occurring on other links in the network. Figure 3.8 shows the critical links in Calgary road network as portrayed in the vulnerability analysis.

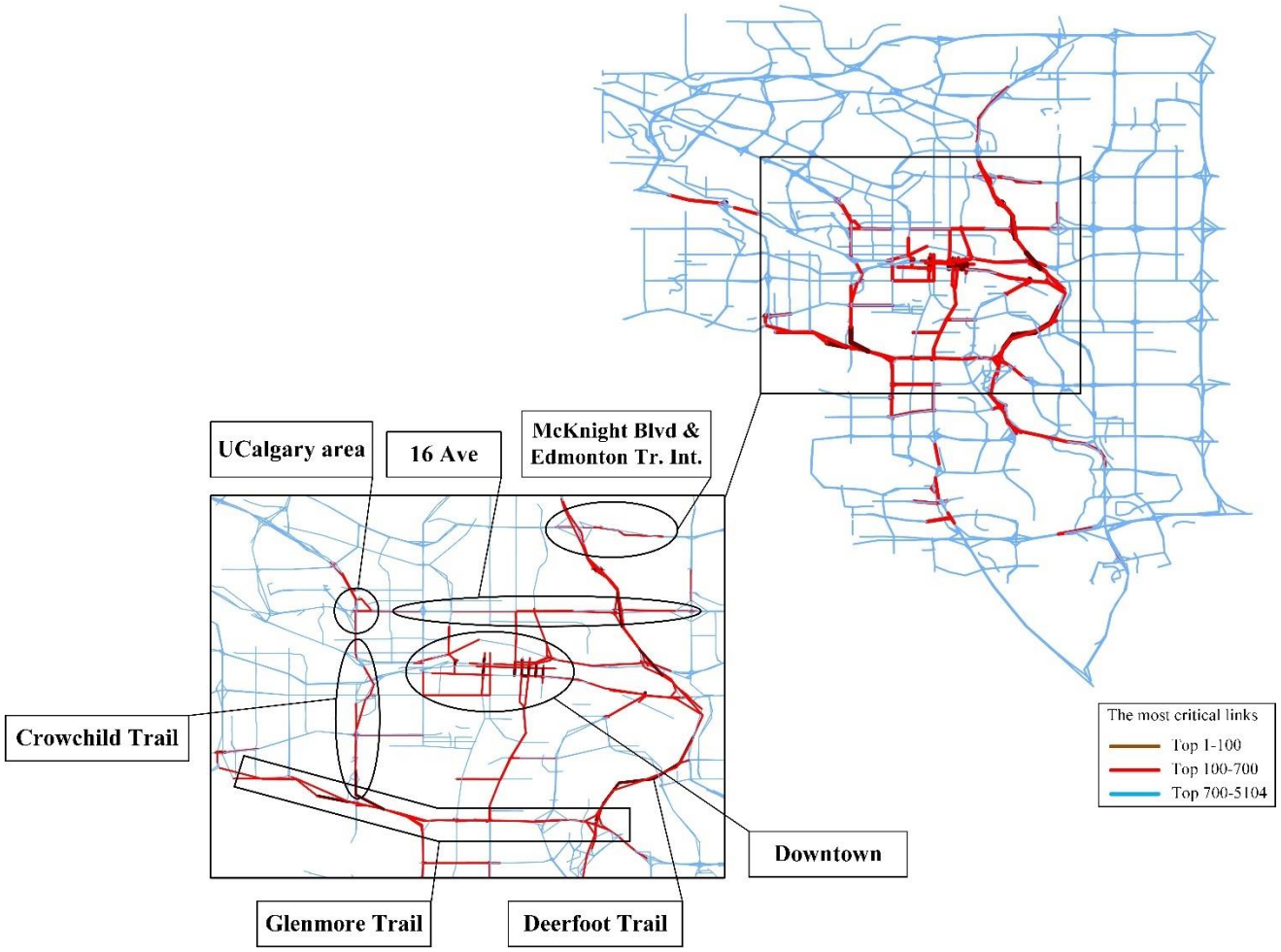

Figure 3.8. The first 800 vulnerable links in the Calgary road network

Notable areas consisting of clusters of critical links are highlighted in Figure 3.8. The most critical links in the Calgary road network are located in the downtown area. Calgary downtown is one of the densest downtown areas of its size in North America and also ranked second in the largest concentration of head offices in Canada (Calgary Economic Development, 2018). High concentration of businesses and high population density in Calgary downtown area result in high travel demand, which can be considered as one of the most important factors contributing to the vulnerability of the links located in this area. In addition, most of the links in Calgary downtown area are one-way streets (e.g., 4th, 5th, 6th,

and 9th Ave, and 4th, 5th, and 9th St SW) with a limited number of turning movements to access these links where most of these alternative roads are themselves one-way road segments (i.e., one might need to use three one-way streets instead of making a left turn). Thus, despite being a dense urban area, measured by the number of links per unit area, Calgary downtown suffers from relatively low connectivity. This is the same case for any other downtown with a large number of one-way streets. Because of the low connectivity between the road segments within the Calgary downtown area, the high network density cannot easily absorb and thus mitigate the congestion caused by disruptive events, which makes this area highly vulnerable to the occurrence of incidents. Another influential factor is lack of cyclicity in the Calgary downtown area. Due to the high number of one-way streets, a few random walks lead to a cycle back to a previously visited node in the downtown area. For instance, visiting the intersection of 5 Ave SW and 5 St SW requires a minimum walk of the length of 8 (i.e., visiting at minimum 8 intersections prior to visiting the mentioned intersection again). Such lack of cyclicity further affects the capability of the downtown area to recover from the consequences of link failure. This finding is consistent with the finding in Zhang et al. (2015), who reported cyclicity to be positively correlated with resilience.

Most of the corridors reported as vulnerable road segments in Chapter 2 are indeed also identified as vulnerable corridors using Equation (3.12) as the measure of vulnerability. Examples are the stretch of the 16 Ave between Deerfoot Trail and 14th St, Banff Trail, the intersection of McKnight Blvd and Deerfoot Trail/Edmonton Trail and the area in the vicinity of the University of Calgary.

In addition, several new corridors are added to the set of vulnerable links by applying the methodology developed in this chapter. Notable examples are Deerfoot Trail, Crowchild

Trial and Glenmore Trail. Deerfoot Trail and Glenmore Trail are Alberta's busiest highways with peak daily congestion lasting as many as four hours (Alberta Transportation, 2018). The high traffic volume on these two corridors is a contributing factor to an average 2,000 collisions annually (Farhan, 2019). In contrast to dense urban areas such as Calgary downtown, where lack of alternative access roads is likely the contributing factor to the vulnerability of links subject to incident, the demand for travel becomes a more contributing factor to vulnerability in areas with lower densities (e.g., major highways such as Deerfoot Trail and Glenmore Trail). The same explanation applies to the part of Crowchild Trail between Glenmore Trail and 32 Ave NW, which connects the lower part of the city to numerous major demand attractors in the City of Calgary such as University of Calgary, Foothills Medical Centre, SAIT and Alberta Children's Hospital. This part of Crowchild Trail contains links with high traffic demand and a very high annual collision rate (around 100 incidents per year) which makes them vulnerable to the presence of disruptive events.

As discussed previously in this section, the majority of the links recognised as vulnerable links in Chapter 2 indeed turned out to be vulnerable network elements by applying the methodology introduced in this chapter. In fact, new corridors are revealed to be also vulnerable that are not identified as vulnerable links by proposed approach in Chapter 2. One possible justification is that in Chapter 2, the vulnerability index is obtained by evaluating the return level of the extreme delays determined in each week as the block interval. The consideration of one-week block interval might underestimate the return level of links with high frequency of collisions (e.g., Deerfoot Trail, Glenmore Trail and Crowchild Trail) where on average more than one incident occurs weekly or even daily. Thus, the approach in Chapter 2 might not completely reveal the vulnerability level of these

links. This issue can be resolved by either choosing smaller block-intervals for links with higher frequency of incidents or applying R^{th} largest order, a method which allows two or more extremes to be chosen in each block interval. The implementation of R^{th} order method could be tricky as it might increase the bias in the obtained extreme database by also including non-extreme travel time observations. Same issue holds for choosing block maxima method with smaller block-intervals especially for links with lower frequency of incidents. Nevertheless, the road vulnerability index developed in this chapter is more holistic as it also considers the impact of all possible incidents on a given link (i.e., both direct and indirect impact of incidents). The consideration of the impact of all incidents is facilitated in this chapter by integrating both the incident and travel time databases. Thus, the vulnerability ranking obtained in this chapter represents more realistically the true vulnerability ranking of links compared to the one obtained in Chapter 2 as is shown in the case study on the Calgary road network.

Comparing my findings on major highways in the City of Calgary with the findings of Chen et al. (2012), my results are consistent, and it can be concluded that corridors with high travel demand in areas with lower urban densities tend to be more vulnerable to disruptive events as compared to links with low travel demand. In addition, it was also found in this chapter that for dense urban areas where the demand of travel is high (e.g., downtown), links with fewer alternative access links and higher frequency of incidents are likely more vulnerable to the presence of disruptive events.

CHAPTER 4: SUMMARY AND CONCLUSIONS

This chapter presents concluding remarks and provides potential directions for new research. Sections 4.1 and 4.2 share an overall research summary and the findings related to Chapters 2 and 3. I suggest areas that may be of interest for future study in Section 4.3.

4.1 Research contributions and findings on the vulnerability analysis by modeling monthly and seasonal extreme travel delay variations

In Chapter 2, I proposed a new vulnerability index by modeling the monthly and seasonal variation in the extreme travel delay.

This research contributes to the previous body of knowledge as follows:

1) *Developing a new class of extreme value distribution called CGEV distribution to describe extreme travel delays observed in a road transportation network*. The proposed CGEV distribution is the product of two GEV distributions used to describe the extreme travel delay variability across two different levels, month-to-month and seasonal. I obtained a closed-form algebraic expression based on a regularized hypergeometric function to represent the PDF and CDF of the CGEV distribution. The proposed CGEV distribution is characterized by its five parameters consisting of two shape parameters (ξ_d, ξ_D), two location parameters (μ_d, μ_D), and the seasonal mean of extreme travel delay (Δ_j). The two shape parameters ξ_d and μ_d characterize the monthly variation in extreme travel delay and the two location parameters ξ_D and μ_D also characterize the monthly variation in extreme travel delay. Finally, Δ_j represents the actual mean of extreme travel delay in a given season *j*.

The proposed CGEV distribution reduces the number of estimated parameters needed to describe the monthly and seasonal variation in extreme travel delay if the variation was intended to be modeled by two separate GEV distributions, i.e., one to describe month-to-month and one for seasonal variability. Moreover, the fact that the parameters of the CGEV distribution characterize the monthly and seasonal variability makes it more convenient to validate the model assumptions compared to that of the GEV distribution because the CGEV distribution relates empirical and analytical CVs for both variability levels.

I calibrated the parameters of the CGEV distribution parameters using actual observed link-based multiyear observed travel time data from the City of Calgary road network. The underlying assumptions used to derive the CGEV distribution are also validated using the same dataset. The results indicate that the mean-SD relationship can be adequately described by a linear function for both monthly and seasonal levels.

2) *Comparing the performance of the proposed CGEV distribution with that of the traditional GEV distribution accounting for both the goodness of fit, i.e., relative log-likelihood value, and model complexity, i.e., number of estimated parameters.* The results indicate that accounting for the seasonality by allowing the same parameters across similar seasons provides a flexible, but not too complex (i.e. in terms of the number of estimated parameters needed), CGEV distribution that is favored over other models that are based on traditional GEV distribution. The results are further validated by comparing the ranking distribution of the developed CGEV distribution across all links in the Calgary road network.

To further validate the assumption that the developed CGEV distribution is able to properly model the overall variation in extreme travel delay, I obtained the seasonal return level of the extreme travel delay, i.e., the level of extreme travel delay expected to be

exceeded on average once every season, for both the CGEV and traditional GEV distributions. In addition, I compared the return level to the actual observed mean of the extreme travel delay across different seasons. I showed the merit of the CGEV distribution in modeling the variability in a large size network and that the CGEV distribution minimizes the error of the actual return level prediction.

The primary improvement of the CGEV distribution over traditional GEV distributions is its ability to model multi-dimensional aspects of extreme travel delay variability and the clear physical interpretation of its parameters in connection with monthly and seasonal variation in extreme travel delay. Modeling the extreme travel delay using the CGEV distribution provides a systematic framework in appraising, assessing, and comparing the extreme travel delay variability, which is crucial to perceive the underlying variabilities in extreme travel delay and their characteristics.

3) *Developing a new data-driven vulnerability approach that accounts for both the stochastic nature of extreme events and link importance*. While the stochastic nature of extreme events is reflected in their return levels, link importance is used as an alternative measure for the relative impact of a link failure on network performance.

I discussed the application of the proposed CGEV distribution to road network vulnerability analysis and proposed a new index of road vulnerability based on the relative importance of links in road networks and their seasonal return levels. I examined the performance of the proposed index in the case of Calgary road network and showed that in dense urban areas, links with lack of access to alternative links are likely more vulnerable to the presence of disruptive events.

4.2 Research contributions and findings on vulnerability analysis by measuring the spatiotemporal impact of incidents

In Chapter 3, I proposed another data-driven vulnerability approach that considers both the consequences of incidents as well as the probability of incidents' impact on the neighbouring links.

Analytical vulnerability analysis usually overlooks important factors that affect the vulnerability of road links. For instance, they model the link closure as a one-time incident and measured the consequences before and after the change with no consideration to dynamic impact of closure during the full-time span from link closure to link recovery. Multi-year observed travel time and incident data could be mined to further unravel these dynamics as these datasets contain important information that reflect the historical spatial and temporal occurrences of link closure and their impacts.

This research contributes to the body of knowledge as follows:

1) *Formulating a novel approach to identify the spatial and temporal impact of incidents; these spatiotemporal impacts of incidents are subsequently used in the vulnerability analysis.* Thus, instead of focusing only on the travel time fluctuation of the link subject to incident, which is the common practice in the literature, the developed approach determines both the time of occurrence and the flow recovery of incidents by monitoring the spatiotemporal propagation of congestion patterns in a collection of links in the vicinity of the link subject to incident (i.e., impact area).

I examined the performance of the developed approach on historical incidents in the Calgary road network. The results indicate that the reported incident duration obtained from start time and clearance time recorded in the dataset is not representative enough of the true

impact of incidents as it overlooks the propagating impact of incidents on the impact area. The proposed approach can determine the enduring spatiotemporal impact of incidents. A more realistic representation of the spatiotemporal propagation of incidents is imperative for a solid vulnerability analysis. The accuracy of such information is crucial in identifying the critical road infrastructure needed to increase road safety and network robustness. The analysis framework that is developed in this research is shown to successfully identify the spatiotemporal propagation of incidents in a road network. In addition, the outcomes of this analysis can be used as a decision support tool to assist transportation policy makers in identifying the critical and vulnerable links in a road network and subsequently prioritize the resources that need to be allocated to improve network wide resilience and robustness.

2) *Developing a new holistic data-driven impact area vulnerability approach that takes into consideration both the probability that links in an impact area are affected by an incident as well as the spatiotemporal propagation of the incident on the impact area*. In addition, I formulated an optimization problem to identify the importance of each link on network performance. The importance index has higher values for links with higher travel paces (e.g., major roads including freeways and arterials) and is lower for links with lower travel pace (e.g., minor links or links with relatively high level of service and operation conditions). The link importance was incorporated into the proposed vulnerability index. I examined the performance of the proposed vulnerability index on the Calgary road network while taking into account the spatiotemporal impact of historical incidents and the importance of links in the Calgary road network. I found that in urban areas with lower urban density, links on corridors with high travel demand are more likely to be vulnerable to disruptive events compared to links with lower demand of travel. In addition, I found that in

dense urban areas (e.g., downtown), lack of connectivity is the most contributing factor to road vulnerability as connectivity can efficiently mitigate the consequences of road closure within the area. The primary feature of the developed vulnerability index is its ability to model multi-dimensional aspects of extreme travel delays caused by incidents in a road network. These multi-dimensional aspects were not examined before.

3) *Estimating the vulnerability ranking of links where both the direct and indirect impacts of incidents are incorporated*. While direct impact considers the impact of all incidents occurring on a given link on the link itself and links in its vicinity, the indirect impact examines the impact of incidents occurring on other links but also affecting the given link. The vulnerability analysis discussed in the literature only measures the direct impact of incidents on a given link, however, network elements in the vicinity of links that are subject to high frequency of incidents are themselves more vulnerable to be affected indirectly by the resulting congestion.

4.3 Future extension

There are numerous possible future directions for this research. For instance, the proposed approach to determine the spatiotemporal impact of incidents in Chapter 3 could be enhanced to identify the impact of other type of incidents other than collisions (e.g., work zones, link closure due to severe weather condition, etc.). The determination of the spatiotemporal impact of non-collision related incidents could be challenging as the impact of these events can last much longer in a road network compared to that of collisions. Consequently, the vulnerability index proposed in Chapter 3 could be further modified to take into consideration the impact of non-collision related incidents.

Another interesting extension of this study is revising the spatiotemporal evaluation approach such that it can be used as an online incident detection technique. The proposed approach in Chapter 3 evaluates the spatiotemporal impact of incidents by analysing the travel pace data during the incident time which requires having access to historical incident pace data. Thus, in the current form, the developed approach cannot be used as an incident detection technique. The provision of real-time traffic information gives transportation authorities the possibility of detecting incidents soon after their occurrences which may help design strategies for traffic management and rerouting information - to be delivered to drivers online through variable-message signs, traffic radio or navigation devices - to mitigate congestion and improve the performance of road networks in the presence of incidents.

Another potential extension of this study is considering possible future advancements in traffic monitoring through connected vehicles. Connected vehicles are becoming information hubs that generate, process, send, and receive vast amounts of data while on the move. I can study how I can make use of introducing new advancements to current traffic data collection methods to improve the accuracy of the proposed spatiotemporal incident evaluation technique. Furthermore, I can study how to harness the massive amount of real-time data provided by the connected vehicle environment to develop more efficient incident detection techniques.

Appendix I

Let's assume that the error term ε_d follows the GEV distribution with location parameter μ_d, shape parameter ξ_d, and scale parameter σ_d. The expected value of ε_d, denoted by $E(\varepsilon_d)$, is equal to 1 as follows:

$$E(\varepsilon_d) = 1 \tag{I-2.1}$$

Replacing the mean value of the error term based on the GEV distribution as $\mu_d + \sigma_d \frac{\Gamma(1-\xi_d)-1}{\xi_d}$ in Equation (I-2.1) results in the following:

$$\mu_d + \sigma_d \frac{\Gamma(1-\xi_d)-1}{\xi_d} = 1 \tag{I-2.2}$$

The scale parameter σ_d is then obtained based on the shape and location parameters by re-organizing Equation (I-2.2) as follows:

$$\sigma_d = \frac{(1-\mu_d)\xi_d}{\Gamma(1-\xi_d)-1} \tag{I-2.3}$$

REFERENCES

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. *In Selected papers of hirotugu akaike*, 199-213. Springer, New York, NY.

Alberta Transportation. (2016). Alberta Traffic Collision Statistics. Retrieved from https://www.transportation.alberta.ca/Content/docType47/Production/AR2016.pdf

Alberta Transportation. (2018). Alberta Highways 1 to 986 - Traffic Volume History 1962-2017. Retrieved April 21, 2020.

Al-Deek, H. & Emam, E. B. (2006). New methodology for estimating reliability in transportation networks with degraded link capacities. *Journal of Intelligent Transportation Systems*, 10, 117-129.

Almotahari, A. & Yazici, M. A. (2019). A link criticality index embedded in the convex combinations solution of user equilibrium traffic assignment. *Transportation Research Part A: Policy and Practice*, 126, 67-82.

Alvarado, E., Sandberg, D.V., & Pickford, S.G. (1998). Modeling large forest fires as extreme events. *Northwest Sci.* 72(Spec. Issue): 66–75.

Amer, A., Roberts, E., Mangar, U., Kraft, W. H., Wanat, J. T., Cusolito, P. C., ... & Zhao, X. (2015). Traffic Incident Management Gap Analysis Primer (No. FHWA-HOP-15-007). United States. Federal Highway Administration. Office of Operations.

Ansari Esfeh, M., Kattan, L., Lam, W. H. K., Ansari, R., Salari, M. (2020). Compound Generalized Extreme Value distribution for modeling the effects of monthly and seasonal variation on the extreme travel delays for vulnerability analysis of road network. Revise and Resubmit in *Transportation research part C: emerging technologies*.

Asakura, Y. & Kashiwadani, M. (1991). Road network reliability caused by daily fluctuation of traffic flow. *In PTRC Summer Annual Meeting*, 19th, 1991, University of Sussex, United Kingdom.

Asakura, Y., Kusakabe, T., Nguyen, L. X., & Ushiki, T. (2017). Incident detection methods using probe vehicles with on-board GPS equipment. *Transportation research part C: emerging technologies,* 81, 330-341.

Åsljung, D., Nilsson, J., & Fredriksson, J. (2016). Comparing collision threat measures for verification of autonomous vehicles using extreme value theory. *IFAC-PapersOnLine*, 49, 57-62.

Åsljung, D., Nilsson, J., & Fredriksson, J. (2017). Using extreme value theory for vehicle level safety validation and implications for autonomous vehicles. *IEEE Transactions on Intelligent Vehicles,* 2, 288-297.

Beguería, S. & Vicente-Serrano, S. M. (2006). Mapping the hazard of extreme rainfall by peaks over threshold extreme value analysis and spatial regression techniques. *Journal of Applied Meteorology and Climatology,* 45, 108-124.

Bell, M. G., Kurauchi, F., Perera, S., & Wong, W. (2017). Investigating transport network vulnerability by capacity weighted spectral analysis. *Transportation Research Part B: Methodological*, 99, 251-266.

Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., & Palutikof, J. (2007). Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change*, 81, 71-95.

Berdica, K. (2002). An introduction to road vulnerability: what has been done, is done and should be done. *Transport Policy*, 9, 117-127.

Bogers, E. A., Viti, F., Hoogendoorn, S. P., & Van Zuylen, H. J. (2006). Valuation of different types of travel time reliability in route choice: Large-scale laboratory experiment. *Transportation Research Record*, 1985, 162-170.

Bordi, I., Fraedrich, K., Petitta, M., & Sutera, A. (2007). Extreme value analysis of wet and dry periods in Sicily. *Theoretical and Applied Climatology*, 87, 61-71.

Brilon, W., Geistefeldt, J., & Regler, M. (2005). Reliability of freeway traffic flow: a stochastic concept of capacity. *In Proceedings of the 16th International Symposium on Transportation and Traffic Theory* (Vol. 125143). Maryland: College Park.

Burke, E. J., Perry, R. H., & Brown, S. J. (2010). An extreme value analysis of UK drought and projections of change in the future. *Journal of Hydrology*, 388, 131-143.

Caers, J., Vynckier, P., Beirlant, J., & Rombouts, L. (1996). Extreme value analysis of diamond-size distributions. *Mathematical Geology*, 28, 25-43.

Calgary Economic Development. (2018). Fact Sheet: Calgary Head Offices. Retrieved from https://calgaryeconomicdevelopment.com/dmsdocument/133

Campbell, K. W. (1982). Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model. *Bulletin of the Seismological Society of America*, 72, 1689-1705.

Castillo, E., Nogal, M., Menéndez, J. M., Sanchez-Cambronero, S., & Jimenez, P. (2012). Stochastic demand dynamic traffic models using generalized beta-Gaussian Bayesian networks. *IEEE Transactions on Intelligent Transportation Systems*, 13, 565-581.

Chakraborty, P., Hegde, C., & Sharma, A. (2019). Data-driven parallelizable traffic incident detection using spatio-temporally denoised robust thresholds. *Transportation research part C: emerging technologies*, 105, 81-99.

Chen, A. & Recker, W. (2000). Considering risk taking behavior in travel time reliability. UCI–ITS–WP–00–24, Institute of Transportation Studies. University of California, Irvine. http://www.its. uci.edu.

Chen, A., Yang, C., Kongsomsaksakul, S., & Lee, M. (2007). Network-based accessibility measures for vulnerability analysis of degradable transportation networks. *Networks and Spatial Economics*, 7, 241-256.

Chen, A., Zhou, Z., & Lam, W. H. K. (2011). Modeling stochastic perception error in the mean-excess traffic equilibrium model. *Transportation Research Part B: Methodological*, 45, 1619-1640.

Chen, B. Y., Lam, W. H. K., Sumalee, A., & Li, Z. L. (2012). Reliable shortest path finding in stochastic networks with spatial correlated link travel times. *International Journal of Geographical Information Science*, 26, 365-386.

Chen, B. Y., Lam, W. H. K., Sumalee, A., Li, Q., & Li, Z. C. (2012). Vulnerability analysis for large-scale and congested road networks with demand uncertainty. *Transportation Research Part A: Policy and Practice*, 46, 501-516.

Chen, C., Skabardonis, A., & Varaiya, P. (2003). Travel-time reliability as a measure of service. *Transportation Research Record: Journal of the Transportation Research Board*, (1855), 74-79.

Cheng, L., AghaKouchak, A., Gilleland, E., & Katz, R. W. (2014). Non-stationary extreme value analysis in a changing climate. *Climatic change*, 127, 353-369.

Cohen, H. & Southworth, F. (1999). On the measurement and valuation of travel time variability due to incidents on freeways. *Journal of Transportation and Statistics*, 2, 123-131.

Coles, S. & Pericchi, L. (2003). Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52, 405-416.

Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). An introduction to statistical modeling of extreme values (Vol. 208, p. 208). London: Springer.

Cooley, D. (2009). Extreme value analysis and the study of climate change. *Climatic Change*, 97, 77-83.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 5, 1583-1606.

D'Este, G. A., & Taylor, M. A. P. (2003). Network vulnerability: an approach to reliability analysis at the level of national strategic transport networks. *The network reliability of transport*, 20, 23-44.

Danczyk, A., Liu, H.X. (2010). Unexpected cause, unexpected effect: empirical observations of twin cities traffic behavior after the I-35W bridge collapse and reopening. *In proceedings of Transportation Research Board*, CD-ROW.

Darayi, M., Barker, K., & Santos, J. R. (2017). Component importance measures for multi-industry vulnerability of a freight transportation network. *Networks and spatial economics*, 17, 1111-1136.

de Jong, G. C. & Bliemer, M. C. (2015). On including travel time reliability of road traffic in appraisal. *Transportation Research Part A: Policy and Practice*, 73, 80-95.

Demšar, U., Špatenková, O., & Virrantaus, K. (2008). Identifying critical locations in a spatial network with graph theory. *Transactions in GIS*, 12, 61-82.

Dial, R. B. (2006). A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration. *Transportation Research Part B: Methodological*, 40, 917-936.

Donovan, B., & Work, D. B. (2017). Empirically quantifying city-scale transportation system resilience to extreme events. *Transportation Research Part C: Emerging Technologies*, 79, 333-346.

Du, Q., Kishi, K., Aiura, N., & Nakatsuji, T. (2014). Transportation Network Vulnerability: Vulnerability Scanning Methodology Applied to Multiple Logistics Transport Networks. *Transportation Research Record: Journal of the Transportation Research Board*, (2410), 96-104.

Duan, Y., & Lu, F. (2014). Robustness of city road networks at different granularities. *Physica A: Statistical Mechanics and its Applications*, 411, 21-34.

Edie, L.C., (1963). Discussion of traffic stream measurements and definitions. *In proceedings of the 2nd International Symposium on the Theory of Traffic Flow*, pp. 139–154.

Engle, R. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics*, 17, 425-446.

Environment Canada (2019). Canadian Climate Normals: 1981-2010 Climate Normals & Averages. Retrieved from http://climate.weather.gc.ca/climate_normals/index_e.html

Erath, A., Birdsall, J., Axhausen, K., & Hajdin, R. (2009). Vulnerability assessment methodology for Swiss road network. *Transportation Research Record: Journal of the Transportation Research Board*, (2137), 118-126.

Ercelebi, S. G. & Toros, H. (2009). Extreme value analysis of Istanbul air pollution data. *CLEAN–Soil, Air, Water,* 37, 122-131.

Esfeh, M. A., Caldera, H. J., Heshami, S., Moshahedi, N., & Wirasinghe, S. C. (2016). The severity of earthquake events–statistical analysis and classification. *International Journal of Urban Sciences*, 1-21.

Farhan, A. (2019). Network-Level safety prediction models for long-range transportation planning. Doctoral dissertation, University of Calgary, Calgary, Canada.

Faturechi, R. & Miller-Hooks, E. (2014). Measuring the performance of transportation infrastructure systems in disasters: A comprehensive review. *Journal of Infrastructure Systems*, 21, 04014025.

Ferrez, J., Davison, A. C., & Rebetez, M. (2011). Extreme temperature analysis under forest cover compared to an open field. *Agricultural and Forest Meteorology*, 151, 992-1001.

Fisher, R. A. & Tippett, L. H. C. (1928). Limiting forms of the frequency distributions of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc*. 24, 180–190

Florida Department of Transpiration (FDOT). (2000). The Florida reliability method: In Florida's mobility performance measures program. Retrieved from http://www.dot.state.fl.us/planning/statistics/mobilitymeasures/reliability.pdf

Fosgerau, M. & Fukuda, D. (2012). Valuing travel time variability: Characteristics of the travel time distribution on an urban road. *Transportation Research Part C: Emerging Technologies*, 24, 83-101.

Fosgerau, M. & Karlström, A. (2010). The value of reliability. *Transportation Research Part B: Methodological*, 44, 38-49.

Frances, F., Salas, J. D., & Boes, D. C. (1994). Flood frequency analysis with systematic and historical or paleoflood data based on the two-parameter general extreme value models. *Water Resources Research*, 30, 1653-1664.

Gao, L., Liu, X., Liu, Y., Wang, P., Deng, M., Zhu, Q., & Li, H. (2019). Measuring road network topology vulnerability by Ricci curvature. *Physica A: Statistical Mechanics and its Applications*, 527, 121071.

García-Palomares, J. C., Gutiérrez, J., Martín, J. C., & Moya-Gómez, B. (2018). An analysis of the Spanish high capacity road network criticality. *Transportation*, 45, 1139-1159.

Gembris, D., Taylor, J. G., & Suter, D. (2007). Evolution of athletic records: statistical effects versus real improvements. *Journal of Applied Statistics*, 34, 529-545.

Gencay, R. & Selcuk, F. (2004). Extreme value theory and Value-at-Risk: Relative performance in emerging markets. *International Journal of Forecasting*, 20, 287-303.

Gilli, M. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, 27, 207-228.

Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une s´erie al´eatoire. *Ann. Math*. 44, 423–453.

Gradshteyn, I. S., & Ryzhik, I. M. (2014). Table of integrals, series, and products. *Academic press*.

Gu, Y., Fu, X., Liu, Z., Xu, X., & Chen, A. (2019). Performance of transportation network under perturbations: Reliability, vulnerability, and resilience. *Transportation Research Part E: Logistics and Transportation Review*, 133, 101809.

Gurobi Optimization, LLC. (2020). Gurobi Optimizer Reference Manual.

Harris, R. I. (1996). Gumbel re-visited-a new look at extreme value statistics applied to wind speeds. *Journal of Wind Engineering and Industrial Aerodynamics*, 59, 1-22.

Hellinga, B. (2011). Defining, Measuring, and Modelling Transportation Network Reliability. Final Report, Delft University of Technology, the Netherlands.

Herman, R., & Lam, T. (1974). Trip time characteristics of journeys to and from work. *Transportation and Traffic Theory*, 6, 57-86.

Holmes, J. D. & Moriarty, W. W. (1999). Application of the generalized Pareto distribution to extreme value analysis in wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics*, 83, 1-10.

Husdal, J. (2005). The vulnerability of road networks in a cost-benefit perspective. *In Proceedings of the Transportation Research Board Annual Meeting (TRB 2005)*, Washington DC, USA (pp. 9-13).

Hyde, T. & Wright, C. C. (1986). Extreme value methods for estimating road traffic capacity. *Transportation Research Part B: Methodological*, 20, 125-138.

Janić, M. (2015). Modelling the resilience, friability and costs of an air transport network affected by a large-scale disruptive event. *Transportation Research Part A: Policy and Practice*, 81, 77-92.

Jenelius, E. & Mattsson, L. G. (2015). Road network vulnerability analysis: Conceptualization, implementation and application. *Computers, Environment and Urban Systems*, 49, 136-147.

Jenelius, E. (2009). Network structure and travel patterns: explaining the geographical disparities of road network vulnerability. *Journal of Transport Geography*, 17, 234-244.

Jenelius, E., Petersen, T., & Mattsson, L. G. (2006). Importance and exposure in road network vulnerability analysis. *Transportation Research Part A: Policy and Practice*, 40, 537-560.

Jeong, Y. S., Castro-Neto, M., Jeong, M. K., & Han, L. D. (2011). A wavelet-based freeway incident detection algorithm with adapting threshold parameters. *Transportation Research Part C: Emerging Technologies*, 19, 1-19.

Ji, Y., & Geroliminis, N. (2012). On the spatial partitioning of urban transportation networks. *Transportation Research Part B: Methodological*, 46(, 1639-1656.

Jiang, R., Lu, Q. C., & Peng, Z. R. (2018). A station-based rail transit network vulnerability measure considering land use dependency. *Journal of Transport Geography*, 66, 10-18.

Jones, E. G., Mahmassani, H. S., Herman, R., & Walton, C. M. (1989). Travel time variability in a commuting corridor: implications for electronic route guidance. Presented at the *First International Conference on Applications of Advanced Technologies in Transportation Engineering*, San Diego, California.

Katz, R. W. (1999). Extreme value theory for precipitation: sensitivity analysis for climate change. *Advances in Water Resources*, 23, 133-139.

Kilpeläinen, M. & Summala, H. (2007). Effects of weather and weather forecasts on driver behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10, 288-299.

Kim, J. & Mahmassani, H. S. (2015). Compound Gamma representation for modeling travel time variability in a traffic network. *Transportation Research Part B: Methodological*, 80, 40-63.

Klophaus, R., & Lordan, O. (2018). Codesharing network vulnerability of global airline alliances. *Transportation Research Part A: Policy and Practice*, 111, 1-10.

Knoop, V. L., Snelder, M., van Zuylen, H. J., & Hoogendoorn, S. P. (2012). Link-level vulnerability indicators for real-world networks. *Transportation Research Part A: Policy and Practice*, 46, 843-854.

Kouwenhoven, M. & Warffemius, P. (2016) Forecasting Travel time reliability in road transport. *ITF Roundtable Reports Quantifying the Socio-economic Benefits of Transport*, *1*, 57.

Kurauchi, F., Uno, N., Sumalee, A., & Seto, Y. (2009). Network evaluation based on connectivity vulnerability. *In Transportation and Traffic Theory 2009: Golden Jubilee* (pp. 637-649). Springer, Boston, MA.

Kütchenhoff, H. & Thamerus, M. (1996). Extreme value analysis of Munich air pollution data. *Environmental and Ecological Statistics*, 3, 127-141.

Kwon, J., Barkley, T., Hranac, R., Petty, K., & Compin, N. (2011). Decomposition of travel time reliability into various sources: incidents, weather, work zones, special events, and base capacity. *Transportation Research Record: Journal of the Transportation Research Board*, 2229, 28-33.

Kwon, J., Coifman, B., & Bickel, P. (2000). Day-to-day travel-time trends and travel-time prediction from loop-detector data. *Transportation Research Record: Journal of the Transportation Research Board*, 1717, 120-129.

Li, M. (2008). Robustness Analysis of Road Networks: a Framework with Combined DTA Models. TU Delft, Delft University of Technology.

Li, S., Li, G., Cheng, Y., & Ran, B. (2020). Urban arterial traffic status detection using cellular data without cellphone GPS information. *Transportation Research Part C: Emerging Technologies*, 114, 446-462.

Li, T., Rong, L., & 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. *Transportation Research Part A: Policy and Practice*, 127, 55-70.

Liu, C., Zhao, M., Sharma, A., & Sarkar, S. (2019). Traffic Dynamics Exploration and Incident Detection Using Spatiotemporal Graphical Modeling. *Journal of Big Data Analytics in Transportation*, 1, 37-55.

Lombardo, F. T., Main, J. A., & Simiu, E. (2009). Automated extraction and classification of thunderstorm and non-thunderstorm wind data for extreme-value analysis. *Journal of Wind Engineering and Industrial Aerodynamics*, 97, 120-131.

Lopez, C., Leclercq, L., Krishnakumari, P., Chiabaut, N., & Van Lint, H. (2017). Revealing the day-to-day regularity of urban congestion patterns with 3D speed maps. *Scientific Reports*, 7, 1-11.

López, F. A., Páez, A., Carrasco, J. A., & Ruminot, N. A. (2017). Vulnerability of nodes under controlled network topology and flow autocorrelation conditions. *Journal of Transport Geography*, 59, 77-87.

Lordan, O., Sallan, J. M., & Simo, P. (2014). Study of the topology and robustness of airline route networks from the complex network approach: a survey and research agenda. *Journal of Transport Geography*, 37, 112-120.

Lu, Q. C. (2018). Modeling network resilience of rail transit under operational incidents. *Transportation Research Part A: Policy and Practice*, 117, 227-237.

Luce, R. D. & Suppes, P. (1965). Preference, utility, and subjective probability. In Luce, R. D., Bush, R., and Galanter, E., editors, *Handbook of Mathematical Psychology*, 3, 249–410. Wiley, New York.

Mahalanobis, P.C., (1936). On the generalized distance in statistics. *Proc. National Institute of Science of India (Calcutta)*, 2, 49–55.

Mahmassani, H. S., Hou, T., & Dong, J. (2012). Characterizing travel time variability in vehicular traffic networks: deriving a robust relation for reliability analysis. *Transportation research record*, 2315, 141-152.

Makkonen, L. (2008). Problems in the extreme value analysis. *Structural Safety*, 30, 405-419.

Matisziw, T. C., & Murray, A. T. (2009). Modeling s–t path availability to support disaster vulnerability assessment of network infrastructure. *Computers & Operations Research*, 36, 16-26.

Mattsson, L. G. & Jenelius, E. (2015). Vulnerability and resilience of transport systems–A discussion of recent research. *Transportation Research Part A: Policy and Practice*, 81, 16-34.

May, A. D., Bonsall, P. W., & Marler, N. W. (1989). Travel time variability of a group of car commuters in north London. Institute of Transport Studies, Univ. of Leeds, Leeds, U.K.

Mayne, J. R. (1979). The estimation of extreme winds. *Journal of Wind Engineering and Industrial Aerodynamics*, 5, 109-137.

Mazloumi, E., Currie, G., & Rose, G. (2009). Using GPS data to gain insight into public transport travel time variability. *Journal of Transportation Engineering*, 136, 623-631.

McFadden, D. (1978). Modeling the choice of residential location. in *Spatial Interaction Theory and Planning Models*, Amsterdam, North-Holland.

McNeil, A. J. & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7, 271-300.

McNeil, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin*, 27, 117-137.

Melchers, R. E., & Beck, A. T. (1999). *Structural reliability analysis and prediction.* John Wiley & Sons.

Minderhoud, M., Botma, H., & Bovy, P. (1997). Assessment of roadway capacity estimation methods. *Transportation Research Record: Journal of the Transportation Research Board*, 1572, 59-67.

Moeini, M. H., Etemad-Shahidi, A., & Chegini, V. (2010). Wave modeling and extreme value analysis off the northern coast of the Persian Gulf. *Applied Ocean Research*, 32, 209-218.

Mohammadi, R. (1997). Journey time variability in the London area. 1. Journey time distribution. *Traffic Engineering & Control*, 38, 250-257.

Mudersbach, C. & Jensen, J. (2010). Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline. *Journal of Flood Risk Management*, 3, 52-62.

Noland, R. B. & Polak, J. W. (2002). Travel time variability: a review of theoretical and empirical issues. *Transport Reviews*, 22, 39-54.

Noland, R. B. & Small, K. A. (1995). Travel Time uncertainty, departure time choice, and the cost of morning commutes, *Transportation Research Record*, 1493, 150-158.

Pan, B., Demiryurek, U., Gupta, C., & Shahabi, C. (2015). Forecasting spatiotemporal impact of traffic incidents for next-generation navigation systems. *Knowledge and Information Systems*, 45, 75-104.

Panagiotakopoulos, D., Majumdar, A., & Ochieng, W. (2009). Characterizing the Distribution of safety occurrences in aviation: an approach using extreme value theory. *Transportation Research Record*, 2106, 129-140.

Papalexiou, S. M. & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49, 187-201.

Polus, A. (1979). A study of travel time and reliability on arterial routes. *Transportation*, 8, 141-151.

Ponzlet, M. (1996). Dynamik der Leistungsfähigkeiten von Autobahnen.Schriftenreihe des Lehrstuhls für Verkehrswesen der Ruhr. Doctoral dissertation, Institute for Transportation and Traffic Engineering, Ruhr University, Bochum, Germany.

Prigogine, I. & Herman, R. (1971). Kinetic theory of vehicular traffic. *Elsevier, New York*.

Pu, W. (2011). Analytic relationships between travel time reliability measures. *Transportation Research Record: Journal of the Transportation Research Board*, 2254, 122-130.

Rämä, P. & Kulmala, R. (2000). Effects of variable message signs for slippery road conditions on driving speed and headways. *Transportation Research Part F: Traffic Psychology and Behaviour*, 3, 85-94.

Renard, B. & Lang, M. (2007). Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. *Advances in Water Resources*, 30, 897-912.

Richardson, A. J. & Taylor, M. A. P. (1978). Travel time variability on commuter journeys. *High Speed Ground Transportation Journal*, 12.

Rivas, D., Caleyo, F., Valor, A., & Hallen, J. M. (2008). Extreme value analysis applied to pitting corrosion experiments in low carbon steel: Comparison of block maxima and peak over threshold approaches. *Corrosion Science,* 50, 3193-3204.

Roach, R. (2010). State of the West 2010: Western Canadian demographic and economic trends. *Canada West Foundation*. pp. 65 & 102.

Rootzén, H. & Tajvidi, N. (1997). Extreme value statistics and windstorm losses: a case study. *Scandinavian Actuarial Journal*, 1997, 70-94.

Rossi, F., Fiorentino, M., & Versace, P. (1984). Two-component extreme value distribution for flood frequency analysis. *Water Resources Research,* 20, 847-856.

Saastamoinen, K. (1993). Effect of road conditions on driving behaviour and properties of the traffic flow. *FinnRa Reports,* 80.

Schlögl, M. & Laaha, G. (2017). Extreme weather exposure identification for road networks–a comparative assessment of statistical methods. *Natural Hazards and Earth System Sciences,* 17, 515-531.

Schnebele, E., & Waters, N. (2014). Road assessment after flood events using non-authoritative data. *Natural Hazards and Earth System Sciences,* 14, 1007.

Sermons, M. W., & Koppelman, F. S. (1996). Use of vehicle positioning data for arterial incident detection. *Transportation Research Part C: Emerging Technologies*, 4, 87-96.

Sethi, V., Bhandari, N., Koppelman, F. S., & Schofer, J. L. (1995). Arterial incident detection using fixed detector and probe vehicle data. *Transportation Research Part C: Emerging Technologies*, 3, 99-112.

Sharma, P., Khare, M., & Chakrabarti, S. P. (1999). Application of extreme value theory for predicting violations of air quality standards for an urban road intersection. *Transportation Research Part D: Transport and Environment*, 4, 201-216.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 367-377.

Strehl, A., & Ghosh, J. (2002). Cluster ensembles---a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research*, 3, 583-617.

Sumalee, A., Pan, T., Zhong, R., Uno, N., & Indra-Payoong, N. (2013). Dynamic stochastic journey time estimation and reliability analysis using stochastic cell transmission model: Algorithm and case studies. *Transportation Research Part C: Emerging Technologies*, 35, 263-285.

Sun, C., Hao, J., Pei, X., Zhang, Z., & Zhang, Y. (2016). A data-driven approach for duration evaluation of accident impacts on urban intersection traffic flow. *In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1354-1359). IEEE.

Susilawati, S., Taylor, M. A., & Somenahalli, S. V. (2013). Distributions of travel time variability on urban roads. *Journal of Advanced Transportation*, 47, 720-736.

Talley, W. K. & Becker, A. J. (1987), On-time Performance and the exponential probability distribution, *Transportation Research Record*, 1198, 22–26.

Tampere, C.M., Stada, J., Immers, B., Peetermans, E. and Organe, K. (2007). Methodology for identifying vulnerable sections in a national road network. *Transportation Research Record,* 2012, pp.1-10.

Tarko, A. P. & Songchitruksa, P. (2005). Estimating the frequency of crashes as extreme traffic events. In *84th Annual Meeting of the Transportation Research Board*,Washington, D.C.

Taylor, M. A. P. (1982). Travel time variability-the case of two public modes. *Transportation Science*, 16, 507-521.

Taylor, M. A. P., & D'Este, G. M. (2004). Safeguarding transport networks: assessment of network vulnerability and development of remedial measures. *Australian Journal of Multidisciplinary Engineering, Special Edition on Engineering a Secure Australia*, 13-22.

Taylor, M. A., Susilawati. (2012). Remoteness and accessibility in the vulnerability analysis of regional road networks. *Transportation research part A: policy and practice*, 46, 761-771.

Taylor, M.A.P. (2017). Vulnerability Analysis for transportation networks. *Elsvier*, Oxford.

Turner, J. K. & Wardrop, J. G. (1951). The variation of journey time in central London. Road Research Laboratory Note. RN/1511/JKT. JGW.

Turner, S.M., Best, M.E., Schrank, D.L., (1996). Measures of effectiveness for major investment studies. Report No. SWUTC/96/467106-1, Texas Transportation Institute, Texas A&M University, College Station, TX.

Van Lint, J. W. C., Van Zuylen, H. J., & Tu, H. (2008). Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transportation Research Part A: Policy and Practice*, 42, 258-277.

Voltes-Dorta, A., Rodríguez-Déniz, H., & Suau-Sanchez, P. (2017). Vulnerability of the European air transport network to major airport closures from the perspective of passenger delays: Ranking the most critical airports. *Transportation Research Part A: Policy and Practice*, 96, 119-145.

Wang, H., & Song, M. (2011). Ckmeans. 1d. dp: optimal k-means clustering in one dimension by dynamic programming. *The R journal*, 3, 29.

Wang, Y., Dong, W., Zhang, L., Chin, D., Papageorgiou, M., Rose, G., & Young, W. (2012). Speed modeling and travel time estimation based on truncated normal and lognormal distributions. *Transportation Research Record: Journal of the Transportation Research Board*, 2315, 66-72.

Watling, D. & Balijepalli, N. C. (2012). A method to assess demand growth vulnerability of travel times on road network links. *Transportation Research Part A: Policy and Practice*, 46, 772-789.

Willems, P., Guillou, A., & Beirlant, J. (2007). Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function. *Journal of Hydrology*, 338, 221-236.

Williams, H. C. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9, 285-344.

Xu, X., Chen, A., Cheng, L., & Lo, H. K. (2014). Modeling distribution tail in network performance assessment: A mean-excess total travel time risk measure and analytical estimation method. *Transportation Research Part B: Methodological*, 66, 32-49.

Yang, F., Yun, M. P., & Yang, X. G. (2014). Travel Time distribution under interrupted flow and application to travel time reliability. *Transportation Research Record: Journal of the Transportation Research Board*, 2466, 114-124.

Yap, M. D., van Oort, N., van Nes, R., & van Arem, B. (2018). Identification and quantification of link vulnerability in multi-level public transport networks: a passenger perspective. *Transportation*, 45, 1161-1180.

Ye, Q. & Kim, H. (2019). Assessing network vulnerability of heavy rail systems with the impact of partial node failures. *Transportation*, 46, 1591-1614.

Yildirimoglu, M., Limniati, Y., & Geroliminis, N. (2015). Investigating empirical implications of hysteresis in day-to-day travel time variability. *Transportation Research Part C: Emerging Technologies*, 55, 340-350.

Yperman, I., & Tampère, C. (2006). Multi-commodity dynamic network loading with kinematic waves and intersection delays. *Proceedings of the DTA*.

Zhang, X., Miller-Hooks, E., & Denny, K. (2015). Assessing the role of network topology in transportation network resilience. *Journal of Transport Geography*, 46, 35-45.

Zheng, L., Ismail, K., & Meng, X. (2014). Freeway safety estimation using extreme value theory approaches: A comparative study. *Accident Analysis & Prevention*, 62, 32-41.

Zhong, R. X., Xie, X. X., Luo, J. C., Pan, T. L., Lam, W. H. K., & Sumalee, A. (2019). Modeling double time-scale travel time processes with application to assessing the resilience of transportation systems. *Transportation Research Part B: Methodological*, 132, 228-248.