**The Dissertation Committee for Jack William Bringardner certifies that this is the approved version of the following dissertation:**

# Application of a Subnetwork Characterization Methodology for Dynamic Traffic Assignment

**Committee:**

Randy Machemehl, Supervisor

Stephen Boyles

Zhanmin Zhang

Junfeng Jiao

William O'Brien

# Application of a Subnetwork Characterization Methodology for Dynamic Traffic Assignment

by

**Jack William Bringardner, B.S.C.E, M.S.E.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**December 2014**

# Dedication

To Victoria Grace Bill, thank you for your love and support.

# Acknowledgements

I would like to thank Dr. Randy Machemehl for his guidance and support during my graduate studies. He and the other faculty and staff at the University of Texas at Austin have been great leaders in my life and have encouraged me to become a better person. I am also grateful for everything Lisa Smith has helped me with over the last few years. The opportunities I had through my collaboration with other researchers made my accomplishments possible. I value the friendship I developed with Dr. Mason Gemar who was a great mentor and partner on several research questions in my graduate career.

I also want to thank my committee: Dr. Stephen Boyles, Dr. Zhanmin Zhang, Dr. Junfeng Jiao, and Dr. William O'Brien for their advice on a wide range of topics from dissertation writing to career development. I owe a special thank you to Dr. Boyles for generously allowing me to be a teaching assistant for his undergraduate course. A long list of other friends and colleagues have helped guide me through graduate school: Matt Pool, Travis Owens, Dr. Natalia Ruiz, Dr. Jen Duthie, Michael Levin, Dr. Jennifer Miller, Dennis Bell, Mubassira Khan, Kris Powledge, David Kan, Anisah Cross, and Nezamuddin. With their insight and opportunities provided through their connection I would not be where I am today. I would also like to thank Dr. Thomas Connolly for the opportunities to develop my teaching skills, and Pam Dahl for the opportunity to be a teaching assistant in the general engineering program. Finally, I would like to thank the sponsors for the fellowships I have received while attending the University of Texas at Austin.

Also, thank you to my family and friends for their support throughout this process.

# Application of a Subnetwork Characterization Methodology for Dynamic Traffic Assignment


Jack William Bringardner, Ph.D.

The University of Texas at Austin, 2014

Supervisor: Randy Machemehl



The focus of this dissertation is a methodology to select an appropriate subnetwork from a large urban transportation network that experiences changes to a small fraction of the whole network. Subnetwork selection techniques are most effective when using a regional dynamic traffic assignment model. The level of detail included in the regional model relieves the user of manually coding subnetwork components because they can be extracted from the full model. This method will reduce the resources necessary for an agency to complete an analysis through time and cost savings. Dynamic traffic assignment also has the powerful capability of determining rerouting due to network changes. However, the major limitation of these new dynamic models is the computational demand of the algorithms, which inhibit use of full regional models for comparing multiple scenarios. By examining a smaller window of the network, where impacts are expected to occur, the burden of computer power and time can be overcome. These methods will contribute to the accuracy of dynamic transportation systems analysis, increase the tractability of these advanced traffic models, and help implement new modeling techniques previously limited by network size. The following describes how to best understand the effects of reducing a network to a subarea and how this technique may be implemented in practice.

## Table of Contents

## List of Tables

## List of Figures

# SECTION 1: ORIGIN

This dissertation is divided into three major sections to emphasize the importance of the origin, route, and destination assumptions behind traffic assignment. These elements and how they are modeled are a fundamental limitation in traffic simulation. The aggregation of trip ends to zones and algorithms that predict vehicle routing influence the behavior of traffic models. These are the structures that traffic engineers and planners assume to be most effective. Since these are the standard assumptions for analyzing transportation systems, users must keep them in mind when developing new methods. Although the overall structure in this paper is unique, it will still follow the standard dissertation format. The origin focuses on the root of the problem through an introduction and literature review, the route describes the methodology and research design, and the destination focuses on a discussion of the analysis and conclusions.

The origin section will introduce the state of the art in traffic network modeling. Network assignment plays a large role in the current need for effective multimodal dynamic transportation modeling. These models are necessary to make informed decisions as developments are made to the transportation system. Several issues need to be addressed including multi-resolution modeling, geospatial data requirements, and model components. A survey of the appropriate literature will provide context for the role and contribution of these research efforts.

# Chapter 1: Introduction

Networks are ubiquitous in modern civilization. Telecommunications, electrical transmission, air travel, and city development all function with similar network principles. The size of the network that is of interest for investigation is dependent on the analysis questions and the operations of the network type. A global network represents the largest extents of the system and may be of interest for large scale impacts, but when a change to the network is minor it is often not necessary to consider the entire network for operational analysis. Intuition would indicate that the area of the network of interest for analysis is related to the magnitude of the proposed perturbation. Depending on the alteration in the network structure the analysis extents could be reduced to a nation, state, city, or even smaller. This concept is most relevant in a robust network such as an urban roadway system. There are numerous applications for examining smaller portions of a transportation system such as incident management, construction work zones, or adding lanes to a roadway.

## 1.1 Background

Dynamic traffic assignment is an emerging regional traffic network simulation technology that is being investigated by transportation planners and engineers alike. A DTA model's most powerful capabilities is predicting impacts of changes to a transportation network, particularly in locations where rerouting is anticipated (Hardy, 2009). For instance, work zones in a downtown area can be modeled to evaluate what impact they may have on local traffic. This type of analysis could be used to help prepare and evaluate traffic control plans (TCP) and subsequently mitigate the increase in user cost incurred during proposed construction. But, in order to do this in the most effective manner, identification of the precise difference between base network conditions and those resulting from a network impact is required. It is likely that an analyst would wish to implement multiple scenarios - a task that is lengthy enough without having to address variability in the results. As the Federal Highway Administration (FHWA) traffic analysis toolbox for DTA warns, “The number of model runs and the time it takes to run the model set should be considered in the scope. For a smaller model this may not be significant effort, but in larger models this could take longer and impact the schedule and resources” (Sloboden et al., 2012).

One strategy for efficiently using the power of DTA while managing the computational requirements is subnetwork analysis. Rather than simulating an entire regional network, a modeler can investigate a subarea around a network modification. Current practices involve using jurisdictional or physical boundaries to select a subarea. However, these subnetwork procedures, also known as windowing because of the limited focus relative to the whole picture, often rely on engineering judgment with no rigorous defense of the window implemented. The detailed analysis described in this dissertation uses statistical techniques to compare the efficiency of different subareas used for an impact analysis. The major concepts behind these subnetwork recommendations are presented as methods for comparison, prediction, and implementation.

Static traffic assignment (STA) has been the primary traffic modeling tool used by metropolitan planning organizations (MPO) for long range planning. Most MPOs continue to use STA because their datasets and travel demand procedures have been designed around it. There is a strong motivation to move to DTA because it is a better model for predicting more realistic results. A few MPOs are working on the transition to DTA much like the transition from trip-based travel demand to activity-based travel demand. The major problem for this transition is the input requirements for these advanced models (requiring extensive data collection) and the potential issues in coordinating activity-based travel demand and dynamic traffic assignment. Historically, iterative methods for generating trip-based origins, destinations, modes, and routes have been implemented in practice. A standardized procedure has not been finalized for combining activity based modeling and DTA although some issues have been addressed (Lin et al., 2008). However, based on the current incorporation of DTA in practice it is likely to become the dominant traffic assignment method in the future.

The basic structure of dynamic traffic assignment combines elements of microsimulation and STA, the two most common traffic modeling tools. STA uses network optimization algorithms to assign commuters to a route that minimizes their travel time, with the assumption that all users know the available path travel times. Microsimulation uses traffic flow theory to propagate users through the network by tracking the movement of every vehicle. DTA uses a slightly more aggregate traffic flow model than microsimulation, and also incorporates user equilibrium principles from STA.

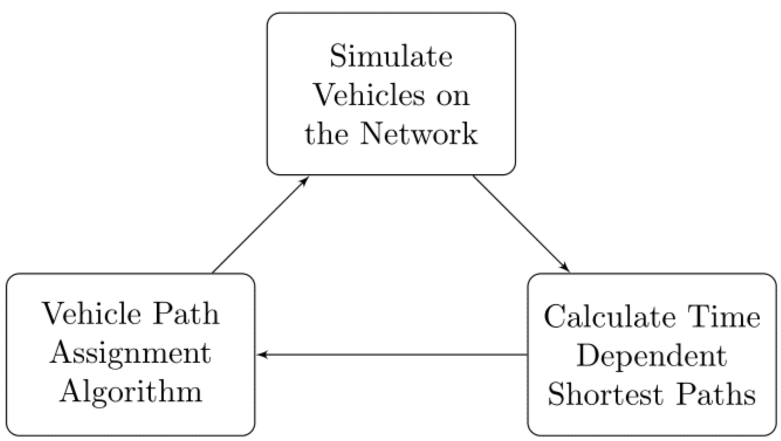



Figure 1: Iterative Process for DTA Algorithms

Figure 1 depicts the iterative process for DTA algorithms to reach convergence. Shortest paths are calculated with respect to time periods of simulation, called time dependent shortest paths (TDSP). Then, a variety of path assignment algorithms can be used to set the route choice for individual vehicles. The traffic flow model in the DTA software then simulates the vehicles using their assigned paths. Each of these steps corresponds to a fundamental principle of a true DTA simulation (Chiu et al., 2010). Figure 2 reveals these relationships to elements in Figure 1. DTA must use experienced travel times, that is, travel times determined from simulation rather than instantaneous travel times like those generated for STA. TDSP algorithms account for this principle. DTA must also incorporate the principle of dynamic user equilibrium (DUE). Like STA user equilibrium (UE), DUE must have all vehicles on equal and minimal travel time paths; but, for DTA UE must hold for each simulation time period. Path assignment algorithms aim to move towards DUE. Finally, DTA must have time varying traffic flow conditions, this principle is what makes it dynamic rather than static. Time varying conditions are incorporated by simulating vehicles with the traffic flow model.

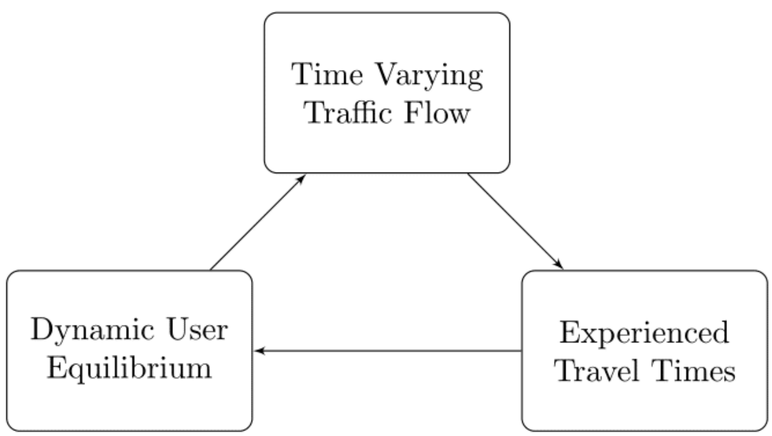



Figure 2: Corresponding Required DTA Components to the Iterative Algorithm

## 1.2 Motivation

As the adaptation of regional DTA models becomes more prevalent in regional planning organizations, the availability of these networks will motivate traffic engineers to utilize this data for traffic simulation. Creation of the network for traffic modeling is one of the most time consuming steps of the traffic modeling process. Not only will engineers have access to high quality network features and attributes (expected to improve as geospatial data standardization efforts continue), but they will also have traffic counts – from DTA results – at all locations in the network. This investigation into the influence of network attributes on results can aid decisions on what area to use for network alteration evaluation; in addition, practitioners may use this information to predict the impacts of reducing a DTA network.

## 1.3 Challenges

Subnetwork selection is not a trivial issue because of the complex nature of dynamic traffic assignment. Understanding the relationship between traffic simulation inputs and outputs is difficult enough without random model components. DTA is not a deterministic model because it can contain stochastic elements to imitate real-world behavior and other random model components. Whether the source of this randomness adds realism or is an artifact of an algorithm, it may be dealt with similarly if the appropriate metrics are chosen. This randomness is difficult to track because of the disaggregate, dynamic nature of the model operations.

DTA model convergence is also an area that has not been fully defined. The cause and effect relationship for DTA processes is obfuscated by considering time dependent operations. Since DTA simulation is influenced by both spatial and temporal factors, randomness may propagate throughout the network in an unpredictable manner. In addition, the potentially large scale of a DTA network provides more opportunity for deviation to occur. For these reasons, users of DTA models require a means to predict and understand the error associated with model results. This error can also be a valuable indicator for what happens when reducing a full network to a subnetwork. A primary need for dealing with the complexity of DTA is controlled data inputs in order to isolate the outcome of the DTA algorithms. This experimental control is best accomplished through implementing standards for geospatial data including link attributes, node characteristics, and demand data.

As transportation modelers improve planning tools the quality of data can be a limiting factor for the accuracy of predictions. DTA models require extra consideration because of their sensitivity to errors in input data. This sensitivity is a response to DTA capturing the temporal changes in traffic flow by generating a time dependent network user equilibrium. DTA incorporates the STA assumption that users choose their shortest path travel time and simulates vehicle operations using principles of traffic flow theory, which give it a level of sophistication that most traffic simulations do not have. The best available DTA algorithms are computationally demanding because they combine elements of regional planning models and microsimulation. In the future, this complexity will increase as even more model components are disaggregated. The geographic information systems (GIS) environment has provided transportation professionals with a means for coordinating and standardizing data collection efforts.

The growth of digital datasets for transportation modeling has been a gift and a curse. Sharing and standardizing information – a primary advantage of geographic databases – has been neglected. Instead, agencies have focused on taking advantage of new computing capabilities without foresight for the potential that this data, in a standard form, has for interoperability. Transportation engineers and planners have been excited by the potential of implementing geodata, but they also need to address data redundancy, security, storage, and integrity. Ultimately, the traffic industry hopes to continue to sharpen its engineering tools with these spatiotemporal datasets – transportation data is of little use without a time and space stamp – to produce more accurate results for simulation and modeling. The greatest tasks related to geospatial tools and data are establishing a data structure that harmonizes the needs of different transportation models, implementing them in professional engineering guidelines, and coordinating the development of experimental algorithms. The latter is of most importance to accomplish the goal of bridging the gap between the state of the art research and practice.

In an ideal world, transportation professionals would follow the same set of data requirements and evaluation platforms for decision making. In reality there exists a wide range of software options and levels of data quality. While this is a great burden, it was also a necessity for the evolution of software tools in traffic engineering. The competition in technological development for simulating transportation operations has taken us from a naïve understanding of an annual average daily traffic with some temporal resolution and growth factors (time of day, day of year, rate of change etc.) to a disaggregate conceptualization of vehicle by vehicle traffic flow. In this process, multiple software packages emerged and created potential issues when we are trying to use computer-aided informatics for allocating resources to projects. This may happen within modeling resolution (e.g. VISSIM and CORSIM) or between modeling resolution (e.g. Synchro, CORSIM, TransCAD, and VISTA). If city A wants to promote project B with results produced by simulation program C and city D wants to promote project E with results produced by simulation program F, then the decisions made will only be informed given that platforms C and F are comparable. Different organizations develop the alterative platforms which makes it difficult to quantify their comparability. A more noble concern such as selecting design alternative X, Y and Z to maximize safety and effectiveness of the transportation network may be derailed by an inappropriate input or a failure to identify anomalies in the mathematical setup. To prevent a negative reputation, practitioners require an impartial means to verify outputs. Such objectivity demands providing access to controlled input data and communicating functionality of the analysis tools.

Great strides have been made in the inventory of roadway data, most centrally by the National States Geographic Information Council (NSGIC) and the Federal Highway Administration (FHWA). These central agencies taking the lead is critical to eliminate redundancy of datasets across city, state, and federal levels of government. In NSGIC's strategic plan, known as Transportation for the Nation (TFTN), the council proposes its vision for a unified geospatial transportation dataset beginning with a need for high quality centerline data (NSGIC, 2011). As the Transportation for the Nation plan moves forward, considerations for geographic information systems (GIS) inventory should extend to different levels of traffic analysis identified by Tamminga, et al. 2013. The Federal Highway Administration efforts to standardize use of modeling tools have been published in the Traffic Analysis Tools program (FHWA, 2014). Coordinating these standards by viewing them as the data inputs and data processing of the same problem could benefit both endeavors as they move forward. As geometric design engineers need accurate field measurements from professional surveyors, traffic engineers are increasingly in need of high quality information – the only difference is that their "field" is often digital, rather than physical.

## 1.4 PROBLEM STATEMENT

DTA models are often created from regional networks and their algorithms require a large number of calculations, which makes running a full network a computationally intensive process. The largest regional models can take days or weeks to converge at an acceptable outcome, a problem that may be overcome by examining a subarea. Even advanced computing strategies like parallel processing have limitations, since certain algorithms, such as the dynamic traffic simulation, are not easily parallelized. Oftentimes, new algorithms are proposed for DTA but are limited to a small test network, like Sioux Falls, due to difficulties associated with large networks. In addition to testing multiple impact scenarios this subnetwork method may help to advance DTA by enabling the use of some innovative breakthroughs.

The aim of this research is to define how measures may be used to determine the appropriate selection of a subnetwork, such that it maintains much of the capability of the full model to predict impacts. Statistical error measures are robust in this application because they can account for the randomness and complexity of the DTA model. Understanding how statistical measures quantify the error between the subnetwork inputs, their time-dependent origin-destination (OD) matrices, can help identify the proper subnetwork selection metrics. Developing a statistical model is ideal for providing the user with evidence to defend the subnetwork used for analysis. This process will help generate effective subnetworks that can allow users to evaluate multiple scenarios efficiently. Ideally, the user could find a balance between the accuracy of the model (large enough to contain an impact's effects) and the time it takes to run the model (small enough to reduce computation time).

Although advances in computer technology will likely reduce the amount of time to run current DTA models, corresponding increases in the complexity of future models and burden on the analyst to provide more detailed results are likely to offset these advances. Research efforts have been and will continue enhancing the detail of DTA traffic flow models (Jin and Boyles, 2014; Yperman, 2007); addressing intersection control (Mitsakis et al., 2011; Paz and Chiu, 2011); disaggregating zones and intelligently locating centroids (Qian and Zhang, 2012); simulating with a shorter time step or continuously simulating (Han et al., 2012; Nezamuddin and Boyles, 2012); interfacing intelligent transportation systems and DTA output (Chiu, 2013; Mahmassani, 2001); and incorporating reliability into vehicle routing (Boyles and Waller, 2011). The use of subnetworks will be an important approach for advanced DTA applications where computational power is limited compared to the scope of data processing.

### 1.5 Contribution

Advancements in traffic modeling provide higher resolution results at the cost of computational demand that has yet to be matched by processing power. In an attempt to more accurately predict real world impacts, modelers will continue to disaggregate inputs and add complexity to the solution algorithms. As these developments continue, consideration must be made for the spatial scale at which the transportation system is being modeled. While it may seem easier to simulate the entire network, resource constraints necessitate a better method. Subnetwork analysis provides an understanding of what inputs may be aggregated to maintain the level of detail desired for the outputs.

The primary contributions of this dissertation are a concise literature review of subnetwork concepts, a procedure for comparing subarea demand inputs, a method for predicting the effect of a network impact on subarea demand inputs, and how to implement these methods in a GIS platform for subnetwork analysis. This study builds on the concepts developed in Mason Gemar's dissertation, Subnetwork Analysis for Dynamic Traffic

Assignment: Methodology and Application (Gemar, 2013). The resulting techniques of these studies will provide a robust toolset for practitioners to begin creating informed subnetworks. Integrating these methods with existing traffic control analysis tools can add an extra level of sophistication. Most traffic microsimulators rely on existing traffic data, limited to a "before case" information, but using DTA results can provide microsimulation with rerouting in the "after case." DTA subarea selection is a critical tool in multi-resolution modeling. The ultimate goal is to enhance the accuracy of traffic simulation analysis and reduce the time required for carrying it out.

### 1.6 Overview

This investigation was also incorporated into a previously created tool for making traffic control calculations in GIS (Bringardner, 2012). By integrating the GIS tool with this subnetwork selection process better decisions can be made based on dynamic data. This technique is ideal for implementing capacity reductions to a network such as in the development of a traffic control plan. Given a particular construction plan an engineer may want to determine what control scenario will limit congestion in the surrounding area. An engineer could use this GIS interface to input the location and magnitude of a proposed traffic control plan and the tool would extract a recommended subnetwork. This information could be communicated to DTA software to create a network model that can be simulated in a fraction of the time it would take for the full model. This process could be repeated for each proposed traffic control scenario and the model outputs could be compared for desirable measures of effectiveness.

Figure 3 displays a flow chart of the overall approach for developing this methodology. Addressing the current state of knowledge, examining the appropriate tools for this type of analysis, and selecting an optimal approach through validation are the key components to this process.

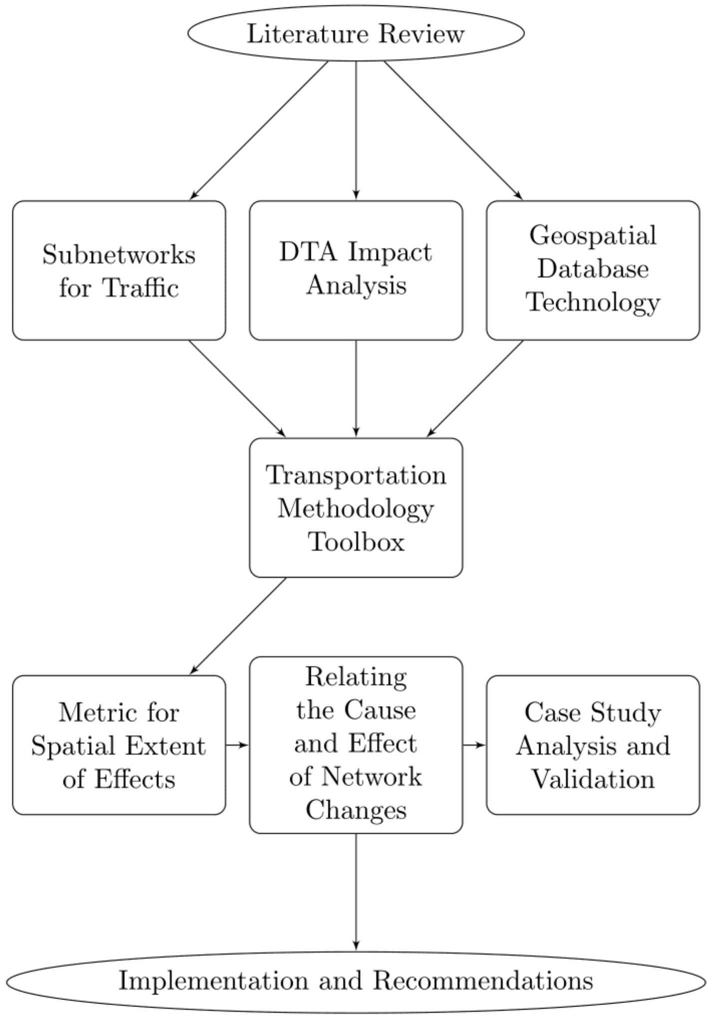



Figure 3: Flow Chart of the Research Approach for this Dissertation

This dissertation will be divided into seven chapters organized into three major sections. The major sections are named to emphasize the importance of the origin, route, and destination assumptions behind traffic assignment. All traffic equilibrium solutions require demand to be moved from an origin zone on the network to a destination zone via the route with the shortest travel time. This framework is both a power and a limitation of this modeling procedure, but is of utmost importance to keep in mind when discussing network assignment capabilities. The first two chapters constitute the origin, or the background of the problem and the current state of the practice. The next two chapters describe the route, or the methodology and proposed solution to the problem. The final chapters define the destination, or validation using a case study analysis, implementation of the solution methods, and conclusions.

**Chapter 1 Introduction** This chapter provides an overview of the background, motivation, challenges, problem statement, and contribution of this study.

**Chapter 2 Literature Review** This chapter will discuss the current body of knowledge associated with the implementation of subnetworks, the use of DTA for traffic impact analysis, the integral role of GIS in advance traffic assignment procedures, and the current analysis techniques for transportation system subareas.

**Chapter 3 Comparison** This chapter describes the statistical comparison techniques for analyzing the impacts of subnetwork size on the accuracy of the DTA model inputs.

**Chapter 4 Prediction** This chapter develops the results of the comparison analysis by allowing the user to predict the effects of their subnetwork size based on the characteristics of a network alteration. It proposes a method for identifying a sufficient subnetwork size relative to the users' unique network configuration.

**Chapter 5 Case Study Analysis** This chapter outlines the final case study data analysis, practical recommendations of these concepts, and geospatial network structure considerations for this procedure.

**Chapter 6 Implementation** This chapter addresses the implementation of the methodology for practical applications. It focuses on the general approach for incorporating traffic analysis tools in GIS.

**Chapter 7 Conclusion** The final chapter summarizes the overall contributions and issues for the methodology.

# Chapter 2: Literature Review

This literature review will elaborate on the concepts in the introduction that can build a foundation for this subnetwork research. Particularly, it will focus on the application based studies involving dynamic traffic assignment (DTA) and geographic information systems (GIS) for transportation. The use of DTA in practice requires some review of the origin of DTA and the major researchers behind its development. There are a number of impact analyses, subnetwork concepts, and GIS traffic modeling practices documented in literature that have inspired elements of this subnetwork study. The goal of the literature review is to provide support for the chosen methodology, and establish the framework for this contribution within DTA development. This dissertation addresses a specific topic within the large realm of DTA, but may have significant implications on the implementation of multi-resolution modeling. While the previous research review will primarily be organized chronologically, certain studies are understood better by connecting them through the contribution of particular research groups.

## 2.1 Impact Analysis using Dynamic Traffic Assignment

Dynamic traffic assignment origins begin with a number of individuals working in the traffic network assignment field. DTA, as it is known today, is primarily simulation based computer software that models regional trips from origins to destinations via time dependent shortest paths. Before simulation algorithms were available, analytical DTA models were used to build the theoretical background for this new form of time based traffic assignment. Simulation allowed for the implementation of DTA for impact analysis related to any network changes – such as incident management and evacuation scenarios.

These early simulation models were worked on by Hani Mahmassani (Mahmassani, 1995), his advisor's advisor Carlos Daganzo (Daganzo, 1995), and Moshe Ben-Akiva

(Ben-Akiva 1998). Mahmassani's effort to create DynaSMART was among one of the first software platforms to provide a fully functioning DTA model. Daganzo developed the traffic flow theory concept of the cell transmission model used in current DTA models. This space and time discretized traffic flow model was constructed around basic traffic flow theory concepts with consideration for the computational demand of DTA algorithms. By dividing links into sublinks, or cells, of equal length, the capacity of the individual cells could be used to propagate traffic on a road network with the added capability of accounting for congestion. Ben-Akiva's DynaMIT was another early software platform which has been used as a study tool for several applications including short-term planning (Sundaram et al., 2011).

The early software initiated the testing of a large number of possible applications for such traffic predictions. Many of these efforts came from the successors of Mahmassani, but the popularity of DTA in recent years has increased the number of researchers in this still relatively small area. It would be difficult to enumerate all people working on DTA development, but some of the major leaders are Ziliaskopoulos, Waller, Chiu, Peeta, and Mouskos. Ziliaskopoulos has developed aspects of DTA such as stochastic modeling and routing problems, and he has documented many of the wide-ranging issues of DTA such as large-scale applications and GIS integration (Peeta and Ziliaskopoulos, 2001; Ziliaskopoulos and Waller, 2000; Ziliaskopoulos et al., 2004). Ziliaskopoulos also helped develop the VISTA (Visual Interactive System for Transportation Algorithms) DTA software with his student Waller. Waller helped bring VISTA (the software used for this project) to implementation ready status and continued investigation of many aspects of DTA. Chiu also went on to develop his own DTA software, DynusT, and has conducted many impact analyses, most notably on predicting the operation of evacuation procedures

(Chiu and Zheng, 2007). Peeta investigated real time deployment of DTA and stability of DTA solutions. Mouskos lead the investigation into several smaller scale impact analyses in the New York area that are more relevant to this dissertation. The efforts of those listed above have sometimes been a collaboration as well as further building off one another's efforts.

Mouskos' efforts are most relevant to this study because he has investigated impacts on different size scales and focused primarily on implementation. One of his early projects, in collaboration with Ziliaskopoulos, was for the New Jersey Department of Transportation using DTA for major corridor analysis (Chien et al., 2005). In the documentation of this project several issues with impact analysis are addressed including data integration such as the geographic road network, origin-destination demand, and model calibration. While the study found DTA to be an effective tool for improving traffic modeling predictions compared to static traffic assignment, one of the greatest obstacles was the lack of a universal data model for the creation of a DTA network. Although the initial coding of a detailed DTA compatible network is a large investment, the model only needs to be updated as permanent changes are made to the roadway inventory. Another major project that Mouskos worked on was an examination of the ability of DTA to predict incident impacts (Sisiopiku et al., 2007; Kamga et al., 2011). This project, which would eventually lead to Kamga's dissertation, quantified the delay and rerouting expected from an incident. The results could then be used to create more informed incident response plans, similar to the potential analysis that could be used for traffic control planning. Building off their previous work they later performed another large scale project for the Alabama Department of Transportation, this time investigating lane reversal (Sisiopiku et al., 2010).

These efforts provided the groundwork for DTA to address problems associated with analyzing multiple scenarios.

The majority of the investigations discussed thus far are from academic projects. Private and research consulting firms have also contributed to the use of DTA in impact analysis. In the 2000s, DTA adaptation began catching on and consulting firms such as Parsons Brinckerhoff were beginning to get involved (Hicks, 2006). INRO, a Canadian based transportation software developer, produced the leading private commercial DTA software, Dynameq. They have documented calibration and application techniques of DTA including comparing results to observed data and implementing improved traffic assignment algorithms (Mahut et al., 2004; Florian et al., 2008). Regional impacts were analyzed by the Texas Transportation Institute to determine rerouting around planned construction (Pesti et al., 2010). More recently, the North Carolina Department of Transportation generated a work zone analysis and impact statement using DTA to supplement traditional microsimulation analysis (Schroeder et al., 2014). These efforts represent the infiltration that DTA models have had in the market for traffic simulation.

The history of dynamic traffic assignment at the University of Texas at Austin continued after Mahmassani's tenure through Waller, a student of Ziliaskopoulos. Waller's student, Boyles, furthered the investigation into many aspects of DTA including the potentials of multi-resolution modeling. Two of his recent students, Matt Pool and Chris Melson, have studied enhancing the applications of DTA by incorporating it into current planning models, transit planning, and improving convergence (Pool, 2013; Melson, 2013). The University of Texas has strong support from the Network Modeling Center (NMC) at Center for Transportation Research (CTR) and its leaders, Duthie and Ruiz. The NMC has worked on several projects to encourage utilizing DTA in practice. A recent technical report published through their efforts identified issues with the use of DTA for bottleneck analysis (Waller et al., 2013). Their analysis involved the selection of a software platform, performance metrics, model calibration and validation, and data requirements. It provides a good summary of the practical aspects of DTA that have been investigated by the NMC. They have also used DTA as a supplemental simulation tool to investigate projects in the central Texas area much like the North Carolina DOT.

The implementation of DTA is an ongoing research goal, and all researchers in this field have identified the need for a unified effort. The Network Modeling Committee of the Transportation Research Board helped assemble a group of experts to generate the DTA Primer, which is the standard introductory document to dynamic traffic assignment for people unfamiliar with the topic (Chiu et al., 2010). The purpose of the document is to classify true DTA and standardize certain concepts that have been put forth by different studies including the obligatory summary article written by Peeta and Ziliaskopoulos, 2001. The primer set the definition of a true DTA model to require time varying traffic flows, dynamic user equilibrium, and experienced travel times. However, much of the primer focuses on conceptual aspects of DTA. The FHWA traffic modeling toolbox has provided a guidebook for practical implementation.

The FHWA Guidebook on the Utilization of Dynamic Traffic Assignment aims to provide practitioners with best practices for using DTA models (Sloboden et al., 2012). One of the first considerations the guidebook mentions is the scope of the analysis. This includes the time scale, the modeling scale, and the spatial scale. With respect to time the guidebook classifies long term, interim, near term, and real time frames. For this dissertation, it is assumed that the demand is available for the appropriate project time period. Implementing results with different modeling resolutions is a tradeoff between the detail of the traffic simulation and the desired model accuracy. DTA is considered mesoscopic simulation and DTA results may be exported to a more disaggregate traffic model. For those reasons the spatial scale, or network size, is the focus of this dissertation. Rather than specifying an arbitrary network size of a region, subregion, or corridor this analysis will specify the real geographic limits of the traffic impacts created by a network alteration.

The guidebook makes recommendations for geographic limits and project planning. The primary guideline is that the physical extents of the subnetwork include possible alternate routes. Although, the rules presented are mostly ad hoc – suggesting the use of natural barriers and the creation of subarea boundary origins where traffic counts are available. Ideally, the guidebook recommends queuing and congestion due to the scenario should be contained within the subarea. These insights indicate that focusing on rerouting created by congestion may be the best method for subnetwork analysis. The guidebook also includes a number of parameters to establish before investigating: study area, type of facilities, travel modes, management strategy, traveler responses, performance metrics, and operational characteristics. One description of the goal of this study is to supplement the recommendations regarding geographic scope in this guidebook.

### 2.2 Subnetwork Analysis for Dynamic Traffic Assignment

Subnetwork analysis has been around as long as traffic modeling. Any microsimulation model that is built is technically a subnetwork. They were usually built as a standalone network because no larger network existed from which to extract a subarea. However, the concept has been defined prior to any understanding of what influence they had on model results. As early as 1976, subnetworks were recommended for urban traffic networks as a means to reduce the size of the optimization problem (Gartner et al., 1976).

Subarea selection for networks is not only a transportation problem, it has also been used to simplify analysis for telecommunications (Galaand and Scotton, 1996). However, the most useful research on procedures for selecting the appropriate subarea came much later.

Many existing research efforts assume that the subnetwork has been predetermined. Once an appropriate subarea is selected the network elements must be extracted including links, nodes, connectors, and centroids. After the network data has been isolated, several techniques have been proposed for building the subarea origin-destination matrix. The most relevant is what has come to be known as an induced origin-destination matrix (Larsson et al., 2001). This method extracts the vehicle routes from a full model run and identifies the new subnetwork boundary origin. Induced subarea demand has become the framework to expand upon for subnetwork demand creation techniques. Initially, this demand extraction was used for static traffic assignment, but it can easily be applied to DTA by assigning a time period to the subnetwork boundary origin.

A primary application of subnetwork creation is the prospect of multi-resolution modeling. This idea of hybrid mesoscopic-microscopic simulation is popular because it allows for complete coverage of data results, with higher quality data in the areas that are most important (Burghout et al., 2005). This technique also allows the mesoscopic DTA model to make predictions about rerouting or future demand to inform the microscopic simulator of the expected impacts. Hybrid modeling commonly uses the induced subarea demand to more accurately depict vehicle trajectories based on the known distribution of vehicle locations from DTA (Ni, 2011).

Xuesong Zhou was one of the first to detail many of the aspects of subarea DTA analysis (Zhou et al., 2006). Working with Mahmassani, their major focus was on dynamic origin-destination demand estimation. They developed an algorithm for updating the induced traffic assignment route output used to create an OD demand matrix for subnetworks with real world traffic counts. Despite a large number of STA subarea demand estimation studies, implementation of subnetwork demand updates has been shown to provide limited modeling improvements for DTA relative to the effort needed for the updates (Gemar et al., 2014). Zhou, nonetheless, identified a classification of subarea OD pairs: Internal-Internal, External-Internal, Internal-External, and External-External. The external portion of the network was defined as the complement network, and a dummy link that simplifies portions of the complement network is denoted a virtual link. These definitions are important when discussing subnetwork concepts. Although they assumed the subarea boundaries to be known for their study, they identify that further work is required for specifying the appropriate subarea. They also address the common need for subarea analysis when analyzing a large number of scenarios and the importance of data structures to limit the building of analysis networks.

A different approach to identifying the portion of a network impacted by a network alteration has been addressed by several transportation planners. The emphasis for these researches offers an alternative to traditional engineering concepts like flow, speed and density. Instead, they suggest looking at principles like accessibility and vulnerability. One approach was to determine the importance of links in a network by correlating the proportion of OD pairs that use each link to the increase in travel time experienced by those OD pairs (Jenelius et al., 2006). In a study by Knoop et al. (2007) various link robustness indicators were evaluated for their effectiveness in predicting the scale of a link's influence on network congestion. Accessibility and network structure were also analyzed to predict the connectivity value of a link added to the overall network (Chen et al., 2007; Taylor, 2008; Jenelius, 2009). The problem is that these studies offer information about the vulnerability of a link for any given scenario, but do not predict the extents of a particular impact. However, these indicators of robustness may be used to understand the factors behind the spread of network queues, since a more robust system is likely to dissipate traffic congestion in a more contained area.

In recent years, many subnetwork concepts have been developed, but have typically been uncoordinated. Developing a method for dynamic OD estimation from static traffic assignment results reveals the importance of tracking path flows into the subnetwork (Choi et al., 2009). Another dynamic OD estimation method based on static traffic assignment results used an algorithm similar to the one used by Zhou using traffic counts (Xie et al., 2010). A return to the planning concepts translated the network robustness indicators into a capacity-disruption value, something more intuitive for application with traffic engineering models (Sullivan et al., 2010). An investigation of the reach of traffic impacts due to the collapse of a bridge identifies the measurable changes at cordon count lines surrounding the bridge's location at different radii (Danczyk et al., 2010). Yet another different approach attempted to use coordinated signal systems as a means to define a unified subarea (Li et al., 2010). The problem with so many unique efforts is that their application may be limited to a particular type of problem. However, if a complete network model is used to produce induced subarea demand, it is possible to capture capacity or volume disruption at cordon counts around an impact. By accounting for flow change experienced at a chosen distance away and accounting for network features like signals or corridors, a unified subarea method could be developed. Combining elements of the efforts of Sullivan, Danczyk, and Li may help to create a new subarea technique.

The most recent efforts have made advances that are valuable to the future of subnetwork analysis. A major contribution is a conversion of the vulnerability analysis into the concept of an impact area (Chen, 2012). This study introduced the concept of an impact area size parameter representing the number of connected links that extend out from the altered link. In this dissertation, this concept, referred to as "connected order" has been adopted as the primary method of identifying a subnetwork area. Although a simple distance or radius concept was initially considered, preliminary testing found the connected order to be a more robust descriptor. It allows for an understanding of what happens to the flow at the boundary as the subarea incrementally increases, and equally grows the subnetwork in all directions.

The subnetwork analysis process has yet to be fully matured. More specific tools that were used for this research project will be elaborated in section 2.4. The goal is to provide an ideal subnetwork so that the appropriate procedures for estimating dynamic OD demand can be most effective (Deng and Cheng, 2013). Despite the lack of information around the effects of subnetwork aggregation, subareas are a necessary process of most DTA investigations (Binkowski and Hicks, 2013; Hadi et al., 2013). The network modeling teams for any project that uses a subnetwork depends on engineering judgment to produce the appropriate subarea. By providing systematic guidelines, this process may be streamlined and automated to reduce the effort required and improve the appropriateness for such an analysis.

### 2.3 Geographic Information Systems and Advanced Traffic Models

Geographic information systems for transportation (GIS-T) have been used for a few decades, but applications have generally incorporated few of the GIS capabilities. GIS-T has evolved from a data inventory framework to a means of operating on geographic features and their attributes. Traffic engineers are often called upon to make quick planning decisions with potentially wide-ranging impacts on travel time (such as lane(s) availability during maintenance or construction activities). In order to make these decisions, they are relying on disparate, non-integrated data sources some of which were originally conceived during the era of mainframe computers with terminal access. This time-consuming process can benefit tremendously from automated procedures coupled with the commonality of the GIS platform which has emerged as the platform of choice for transportation planning and visualization.

Plans for GIS integration should be designed to harvest its true potential by considering horizontal as well as vertical integration. Transportation for the Nation (a plan to produce a road center-line data set for the entire country) has addressed vertical integration from the federal, state and local agencies perspectives. Horizontally, it has been more difficult to assess the needs of the numerous industries that depend on transportation data inputs and outputs. Within each of the following realms a different level of detail and focus is desired: pavement management, energy, environment, operations research, city and regional planning, geography, travel demand, public transit, traffic engineering, geometric design, site planning, and economic policy (Boxill, 2005). Each of these addresses a component of the same transportation system and these efforts may be coordinated by providing a common language for communication. A broad spectrum of issues from safety analysis to DTA data management can be managed through a well-designed transportation GIS framework (Scopatz et al., 2013). New methods can be implemented in practice faster if redundancy is reduced and integrity of shared data resources is increased (Quiroga and Koncz, 2008; Quiroga et al., 2009).

The key for allowing GIS tools and different transportation models to communicate is a database structure (Hossack and Ortega, 2012). All major features, links and nodes for transportation data, should be labeled with a unique identifier (UID). This allows for database operations to be used for editing data and extracting information. GIS provides an intuitive graphical user interface (GUI) for accessing and changing information without the need for users to remember UIDs. Geographic locations known from familiarity with the project or geocoding of street names can be used to find the feature of interest, and selecting the feature in the GUI corresponds to accessing the database with the UID. These IDs must remain stable as changes are made to the planning network in order to maintain consistency.

When relating two databases or interfacing tools with the dataset, the UID should be used as a variable input. This allows for the function to be performed not only on other features in the dataset, but also the same features of a new dataset. GIS adds another level of flexibility by allowing spatial operations to be performed on features based on their geographic attributes. Rather than hardcoding the ID of a feature or geographic coordinates they should be left as an input that can be altered the next time the code is used. Then spatial relationships between the input feature and neighboring features can be determined using GIS analysis tools. The GIS GUI can also be used to visualize changes in network datasets.

Many GIS transportation applications for planning tools and model integration have been documented (Wang, 2005; Carsjens and Ligtenberg, 2007; Khalesian et al., 2009). A major misconception is the need for traffic simulation to be performed within GIS software. Various traffic modeling and analysis packages, many of which have a GIS interface, have been produced by public and private agencies. The leading traffic simulation software has capabilities that surpass the network analysis included in the standard GIS package. GIS software is often produced for a multitude of applications that are not intended to keep up with the latest microsimulation techniques, traffic flow theory, and dynamic routing capabilities. Analysis based on elementary GIS tool estimations of an origin-destination (OD) cost matrix or shortest path identification should be scrutinized. These tools can be useful for preliminary assessments and it is expected that more transportation models will be incorporated in GIS, but the greatest contribution is the interactive data structure that can coordinate with higher level simulation software (Evangelidis et al., 2005). As transportation models and GIS develop, the database connection between them should be considered their common language.

The benefits that GIS brings to DTA modeling are greater than the ability to edit and organize geodatabases. GIS platforms that are available in DTA software often do not have cartographic capabilities. They may have a general coordinate system, but they will not be able to generate projected data necessary for accurate spatial analysis. Projected spherical coordinates onto a Cartesian plane will allow for accurate distance calculations required to determine spatial relationships. GIS provides a handy tool for creation of a subnetwork or using spatial relationships for making adjustment to subarea demand. Flexibility and customizability provided through GIS can enable emerging subnetwork procedures.

Most DTA software has some capability of geographically displaying the network, which is required for generating a subnetwork. If there is no GUI for interacting with the data, then network connectivity is described using intersection node numbers. This type of coding will make the network difficult to update because attributes like link lengths will have to be hardcoded and manually changed. For ease of use a GIS interface and data formats are the best way to verify and import new networks. The network can be viewed as an interchangeable part or input to a DTA model. A new network or an updated nationally organized dataset such as the Transportation for the Nation concept can replace the existing configuration.

## 2.4 Subnetwork Issues and Origin-Destination Analysis

When planning a roadway construction work zone, the primary consideration is safety and effectiveness of the traffic control scheme. Limiting congestion around the construction area is the key to achieving both of these goals. Traffic simulation is the best way to predict how people will choose to reroute to avoid construction caused by lane closures. Some microsimulation models can perform user equilibrium assignments that could describe traveler path choice changes due to construction, but this capability requires the availability of traffic demand in origin-destination matrix form. Even though such data may be available for the whole urban area, microsimulation of an entire urban area is not feasible due to the network coding effort and compute cycle requirements. Deriving an appropriate subnetwork trip table for microsimulation is the essence of the problem that this dissertation will address. The area-wide origin-destination matrix or "trip table" is a basic input to a dynamic traffic assignment model and the DTA is designed to provide area-wide traffic assignments, so DTA is the ideal tool for making traffic control predictions. However, DTA models created for regional analysis provide detailed link volume information outside the construction area zone of influence at the cost of extensive computation time. Creating a method for identifying this zone of influence (likely a fraction of the whole urban area) and the appropriate trip table could enable much more efficient DTA or microsimulation support for traffic control planning.

Traffic impacts of construction related lane closures are often assumed to be localized to the area near the network modification. Closing down a lane for construction can be represented in a network as a capacity reduction. The Highway Capacity Manual

(HCM) procedure for determining roadway capacity includes the number of lanes as a parameter. However, in practice the capacity reduction is commonly applied as a percent reduction corresponding to the fraction of lanes closed. If traffic demand exceeds the construction site capacity, the capacity reduction is represented in the DTA model as the road segment "filling up" with vehicles which limits the volume of upstream traffic entering the link. Travel times for vehicles on a congested construction route will correspondingly increase. When that route travel time exceeds that of alternative routes, vehicles using it get assigned a different route in the process of moving toward user equilibrium.

Traffic simulation relies on GIS network structures and most simulation methods use random number generators to attempt to capture the stochastic nature of traffic operations (Gartner et al., 2005). Typically, allowing the model to approximate the stochasticity of the real world is desirable. However, when model output is analyzed to determine the true effects of an impact, it becomes difficult to discern the variation associated with random processes from the systematic impacts. Other elements in DTA models have the potential to add randomness without adding realism and can have inconsistent effects on outputs. This study also aims to develop statistical tools robust enough to capture all randomness and isolate the true impacts of a network modification.

If a subnetwork and a full network receive statistically similar demands, resulting link traffic volumes should be similar. The DTA algorithm applied for the subarea and full network scenarios are identical and primary differences should be traceable to the subnetwork boundary. Inside the subnetwork boundary, the DTA will produce results with greater convergence (smaller network can be processed through more iterations relatively quickly) and should correspond to the full network predicted volumes if the boundary conditions have been adequately characterized. Characterizing the subnetwork boundary conditions requires a procedure for collapsing the urban area trip table from possibly several million cells to a very small fraction of the original size. Therefore, the key is to provide the subnetwork with appropriate boundary demands, or at least understand the acceptable level of deviation due to the innate variability of the simulation results.

A fundamental component of the subnetwork analysis process is comparison of large trip matrices to determine statistical or practical significance of differences. There are a number of methods that have been introduced for determining differences in OD matrices, primarily for assessing the quality of dynamic OD tables relative to activity travel surveys or traffic counts (Cools et al., 2010; Marzano et al., 2008). For this study, they will be reviewed for application to subnetwork OD matrices to account for inconsistencies inherent to simulating subnetworks and full networks. Two traditional statistics for measuring differences between OD matrices are the root mean square error (RMSE) and the mean absolute percent error (MAPE) (Djukic et al., 2013). The RMSE and MAPE provide estimates of the variance between the induced subnetwork OD matrices extracted from two different simulations of the same full network. The RMSE is calculated using Equation 1:

$$RMSE = \sqrt{\frac{1}{n}\sum_{\forall i,j,t}(\bar{d}_{ijt} - d_{ijt})^2} \tag{1}$$

*where*

$i = origin$

$j = destination$

$t = time\ period$

$d = subnetwork\ demand\ extracted\ from\ an\ individual\ simulation$

$\bar{d} = average\ subnetwork\ demand\ of\ all\ base\ scenario\ simulations$

$n\ =\ number\ of\ OD\ pairs\ being\ compared$

The mean censored absolute percentage error (MCAPE) is a censored version of MAPE because it prevents any individual absolute percent error from exceeding 100% (Cools et al., 2010). MAPE calculations can produce values over 100% if the difference between the demand and the average demand is greater than the average demand. By using a MCAPE, issues stemming from this result can be avoided. Therefore, this paper proposes incorporating the MCAPE by using Equation 2:

$$MCAPE = \frac{1}{n}\sum_{\forall i,j,t} min\left\{100, \left|\frac{\bar{d}_{ijt}-d_{ijt}}{\bar{d}_{ijt}}\right| * 100\right\} \tag{2}$$

*where the notation is the same as Equation 1*

Both of these measures assume that the simulation runs are independent and come from the same distribution. However, "more often than not, output data from simulation experiments are auto-correlated and non-stationary. This precludes analysis using classical statistical techniques which are based on independent and identically distributed (IID) observations (Gartner et al., 2005)." Despite this violation of assumptions, valuable results have been found from both of these measures and they will be compared to a recently proposed method of measuring the magnitude of differences between OD matrices, the structural similarity (SSIM) index.

The SSIM index was initially proposed as a measure of the differences between two images based on structural degradation (Wang et al., 2004). After gaining acceptance in the image processing field, the SSIM index was later applied to the assessment of OD matrices, with OD pair demands replacing pixel values, to help account for the spatial autocorrelation ignored by traditional statistics (Djukic et al., 2013). In essence*, the SSIM* *index calculates a measure based on the location and variability statistics of the demand values by incorporating the surrounding OD cells*. A major assumption for the SSIM index is that the matrix orientation represents geospatial characteristics, which is addressed in Chapter 3.3.

The SSIM index is capable of capturing spatial effects through the concept of a spatial weights matrix, which is also known as a roving window or convolution kernel (Lillesand et al., 2008). The spatial weights matrix is often a 3x3, 5x5, or 7x7 matrix, but can generally be any odd number by odd number matrix so the central value is unique. The normalized spatial weights are then applied beginning with the upper left corner so that the first central value is surrounded by enough values to match the dimensions of the roving window. The window is then applied to every other OD pair surrounded by enough cells to fit the dimensions of the spatial weights matrix. Due to this restriction, the border cells will not be included in the analysis for the SSIM index. SSIM indices are bounded by one and negative one, where a SSIM value of one indicates the matrices are identical. SSIM values are calculated across an entire OD matrix using the mean of the individual spatial weights matrices, as demonstrated in Equation 3:

$$MSSIM(D,\overline{D}) = \frac{1}{n}\sum\nolimits_{\forall n} \frac{\left(2\mu_{d_n}\mu_{\overline{d}_n}+C_1\right)(2\sigma_{d_n\overline{d}_n}+C_2)}{\left({\mu^2}_{d_n}+{\mu^2}_{\overline{d}_n}+C_1\right)({\sigma^2}_{d_n}+{\sigma^2}_{\overline{d}_n}+C_2)} \quad (3)$$

*where*

$D = \; subnetwork\ OD\ matrix\ from\ an\ individual\ simulation$

$\overline{D} = average\ subnetwork\ OD\ matrix\ of\ all\ base\ simulations$

$n \; = \; number\ of\ times\ the\ spatial\ weights\ matrix\ was\ applied$

$d_n =$

$the\ n^{th}\ entrywise\ matrix\ product\ of\ the\ normalized\ spatial\ weights\ matrix\ \&\ D$

$\bar{d}_n =$

$the\ n^{th}\ entrywise\ matrix\ product\ of\ the\ normalized\ spatial\ weights\ matrix\ \&\ \bar{D}$

$\mu = mean\ of\ d_n, \bar{d}_n$

$\sigma^2 = variance\ of\ d_n, \bar{d}_n; \sigma = covariance\ of\ d_n, \bar{d}_n$

$C_1, C_2 = constants\ to\ prevent\ instability\ of\ SSIM\ (dividing\ by\ zero)$

The goal of these statistics is to test the appropriate metrics for quantifying the sufficiency of subnetwork size. First, the metrics will be tested for their ability to capture changes in the boundary demand as the size of the subnetwork grows. Then, the metrics will be used to gauge the amount of rerouting experienced at the subnetwork boundary relative to the magnitude of capacity reduction. In other words, the proposed metric will be capable of relating the impact size of a traffic control scenario to the accuracy of the demand input for the subnetwork, while accounting for the randomness of the model.

### 2.5 SUMMARY

Bringing together the concepts of DTA impact analysis, subnetworks, GIS and transportation modeling, and the initial procedures for this project helps to lay the foundation for defining an appropriate subnetwork area. The greatest power of DTA impact analysis is the ability to predict rerouting. Recent subnetwork studies have focused on adjusting the subarea boundary demand to capture this rerouting, but a new means of extracting an area that captures the majority of rerouting may make the demand adjustment step unnecessary. Developing a controlled subnetwork experiment requires a standard method for extracting the subnetwork GIS elements from the full DTA model. The database and spatial relationship tools of GIS can accommodate the data setup. After building the subnetwork, the most effective way to analyze changes in the full network is to compare the induced traffic flows across the subnetwork boundary. This requires the examination of the origin-destination matrix for the subnetwork, which can be accomplished using a number of metrics. These error measures can provide a thorough analysis for comparing the impacts of subnetwork sizes, and statistical methods may be used to predict these impacts.

# SECTION 1 SUMMARY: ORIGIN

In transportation planning, the origin represents the traffic analysis zone where trips begin. Typical trip generation begins with identifying the characteristics within the zone that influence the number of trips produced, then using historical data to define those relationships. Similarly, the origin section of this dissertation has provided a summary of the issues surrounding subnetworks and used historical literature to address those issues. In essence, the origin defines the current state of the knowledge and where this dissertation begins. After the origin of the problem has been addressed it important to identify the possible routes that may be used solve it, which are covered in Section 2.

To summarize the focus of this problem, the primary scenario that is addressed in this study, corresponding to the key parameters addressed in the FHWA Guidebook on the Utilization of Dynamic Traffic Assignment introduced in Chapter 2.2, are:

- Study area: Small network
- Type of facilities: Arterials and freeway
- Travel modes: Passenger cars and trucks
- Management strategy: Work zone
- Traveler responses: Pre-trip route diversion
- Performance measures: Volume, travel time
- Operational characteristics: Computer run time

This subnetwork selection methodology contributes to the need for intelligently designed subnetworks rather than ad hoc procedures. The techniques proposed here aim to improve the traffic analysis of networks that requires multiple scenarios. The use of sufficient subnetworks will enable DTA advancements to be implemented more quickly as size is often a prohibitive factor for new algorithms.

# SECTION 2: ROUTE

The route section of this dissertation elaborates on the methodology and the research design for solving the subnetwork problem revealed in the origin section. Vehicle routing in DTA is the process of searching for the minimal travel time path between an origin zone and destination zone for a specified analysis time period. This is accomplished through implementing a shortest path algorithm based on the link costs in the transportation network. These costs can be travel time alone or be generalized to include monetary costs, reliability, and other types of impediments to travel. The origin section described where this study began and the route section will expand on how to get to the destination in the most efficient manner. The goal is to take the lessons learned from previous subarea analysis research and synthesize new methods, with technically rigorous reasoning, for identifying an appropriate subnetwork.

The next two chapters will identify the path set for accomplishing a subnetwork characterization and focus on the most effective methods. These methods were tested using an experimental setup based on representative case studies to verify their applicability. Details of the case study are provided to aid the description of the method. Two effective means for establishing subnetworks, a comparison and prediction technique, are described in detail with reference to their theoretical backgrounds. The comparison method introduces a statistical analysis framework that identifies the difference between two induced subnetwork OD matrices; thus, measuring the difference between full network scenarios at the subarea boundary. Information from the comparison analysis led to a predictive model so that an estimate of the error can be produced without the need to run the full network impact or base scenario. The predictive model will also help to generalize the methodology such that other users may adapt the proposed theory to their unique network.

As a reference for the terms used in this section, the following definitions apply to this subnetwork study:

- Full network: a pre-existing, well calibrated regional model
- Subnetwork: the fraction of the full network used for detailed study
- Complement network: the area of the full network not included in the subnetwork
- Base scenario: the output of the full network model run in its original state
- Impact scenario: the output of the full network model run with an alteration or modification made to some link(s)
- Subnetwork OD matrices: the origin-destination demand input for the subnetwork produced form the induced demand of the base or impact scenario
- Boundary demand: this refers to the flow entering the centroids created at the extents of the subnetwork
- Cold start: beginning the DTA model simulation with no initial path set information
- GIS features: geospatial data types containing information on the geographic location of network components including points used to represent nodes and centroids and lines used to represent links and connectors

# Chapter 3: Comparison

This chapter describes a standard method for creating a DTA subnetwork and a means to estimate the possible error associated with removing part of a network. The methodology that evolved first identifies the differences between the full network and the subnetwork and then compares the induced subnetwork demand matrices for a base scenario and impact scenario. The primary limitation caused by creating a subnetwork is the treatment of trips that originate in the complement network, which is why the boundary demand is of greatest interest. The solution presented in this chapter creates subnetwork trip tables extracted from the base and impact scenario full network and then tests for statistical similarity to determine if the impact scenario effects are detectable at the subarea boundary. Three statistical measures found in the literature are used to compare the subnetwork OD matrices. This chapter will discuss how to compare random subnetwork inputs using these three metrics, how to use statistical tests to compare subnetwork sizes for the three methods, and how these results can be used to build subnetwork recommendations.

## 3.1 Subnetwork Performance Comparison

Imagine dropping a rock in a puddle. The effects are intuitive, waves of disturbed water would radiate from the location of the rock's impact in all directions. Now imagine if the water were channelized by walls similar to a simple rat maze (seen in Figure 4). To complicate things even more, suppose different segments of the maze were filled with liquids of different viscosities. Now the influence of the rock on the surrounding water is not so trivial. The effects on one edge of the maze may be drastically different than the other side of the maze. Using this analogy, the maze represents an abstract form of transportation network and the viscosity of the liquid represents the resistance of the traffic flow to change. If one wishes to predict the effects of the size of the rock, the chaos in the perturbation of the water makes it difficult to use basic kinematic wave principles. Instead, examining the maximum wave height at different points in the maze could provide insight into how the rock's influence spreads. Similarly, DTA results have a behavior that is very sensitive to small changes in condition, which suggest the effects of these changes is best understood by investigating empirical observations.

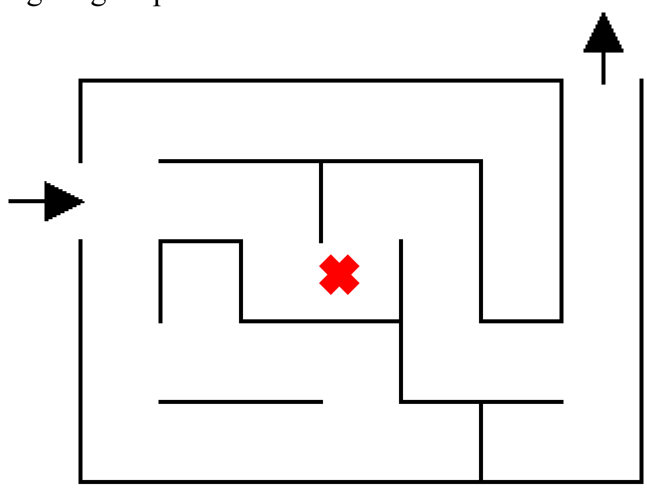


Figure 4: Analogy for the Complexity of Understanding Network Impacts

In the process of developing this methodology for determining subnetwork sufficiency several alternative approaches were evaluated. These approaches were based on standard practices for researching and analyzing transportation problems. A lengthy analysis of link-based statistics for multiple scenarios involved investigating the fundamental traffic parameters: flow, density, and speed. Other metrics were examined as well including travel time and level of service. Upon analyzing the traditional link measures, there was a large amount of variation within multiple simulations of the same scenario. The primary difficulty associated with the randomness of these measures was deriving a summary metric for the entire subnetwork. Comparing different sizes of subnetworks requires an aggregated performance metric, and examining link statistics alone was not robust enough.

An entirely different approach to the subnetwork creation problem was the potential for updating boundary demand. In concept, it is similar to the proposal of Zhou et al. 2006 for estimating subnetwork demand. Rather than using traffic counts, the investigation focused on updating induced boundary demand based on the network model. Initially, an attempt of updating boundary demand with static traffic assignment results was tested. The basic principle was to transform the dynamic network into a static network, using the same demand and links. Then, the base and impact scenario were implemented on the static network through the same capacity reduction. The percent change of flows into and out of the subnetwork were then compared to the results from the DTA model. This method was abandoned after the static traffic assignment results were considered inconsistent with the DTA simulation, the models experienced different magnitudes and locations of changes. While STA shares the principles of user equilibrium, this optimization approach did not improve the DTA model inputs. However, the goal of updating boundary demand was not given up completely.

Another means of updating the subnetwork OD matrix inputs was derived from the destination choice principles used in stochastic network loading (Sheffi, 1984). This is an example where statistical models and optimization methods have been combined in practice. The essence of the discrete destination choice model is using a logit probabilistic model to choose between a set of alternatives with a given utility. In this case, the utility may be represented by the travel times associated with different paths. This concept was applied to trips originating in the complement network. An approximation of the travel times from external regions where used as the utility to select between feasible subnetwork entry point choices. Several configurations of the model were tested, but the algorithms used to extract travel times and estimate the logit model became a limiting factor. The model results recommended adjustments that generated more realistic boundary demand, but the improvements were minor and were offset by the extra time required for running the logit model. Therefore, this was not accepted as a viable method for reducing the time of running the full network model impact scenario.

The methodology was evolving based on the investigation of common tools used for transportation system analysis: standard link-based output metrics, optimization results, and statistical regression models. Insight from testing link-based measures and updating the boundary demand led to a shift of focus toward the interaction of the extents of the subnetwork and the flows across the boundary. This focus inspired the standardization of subnetwork construction. An intuitive method for building the subnetwork is to specify a subarea within a radial distance from the impact location. Using a Euclidian distance does not account for network topology, but it was a valuable first method for testing boundary effects.

Producing analytical measures of the boundary demand involved consideration for the nature of DTA. Since rerouting is a powerful capability of DTA, identifying a method of quantifying rerouting at the subarea boundary became the next goal. After a base and impact scenario have been run, the routes generated in the simulation can be visualized and compared (further discussion of this technique is in Chapter 6.2). It was hypothesized that measuring the distance from the impact to the furthest point of rerouting in the network could be used as the radial distance for the standard subnetwork creation method. Implementing this technique with a few scenarios indicated that it produced too large of a subnetwork. The fraction of the full network included was too large to effectively reduce the time for simulation. Another measure of rerouting was designed to quantify the percentage of displaced vehicles using the impacted link(s) that crossed the boundary. This measure appeared to move the investigation in the right direction, but did not account for the cascading effects of displaced vehicles on other paths – any displaced vehicles use another path causing that path travel time to increase and subsequently rerouting more vehicles away from that path. The method now called for a global measure that summarized the overall impact at the subnetwork boundary.

These steps for building the methodology were critical for the completion of this project. They helped identify the concerns behind what it means for a subnetwork to be sufficient. The issues associated with building an appropriate subnetwork are twofold – the complexity of the model operations and the potential random variation. These problems appear to not be easily captured with an optimization method or new algorithm, and could be addressed more robustly with statistical analysis. This requires a more thorough understanding of the sources of the randomness in DTA outputs.

DTA analyses are desirable because they have been demonstrated to provide more detailed network conditions because of the DTA ability to track temporal changes in demand and traffic, while incorporating user behavior. Within DTA random number generators may introduce variability to replicate real life processes. Examples applications include traffic flow model rounding, loading times for vehicles, creating new shortest paths, OD demand distribution, vehicle class designation, or when vehicles move from one path to another (Peeta and Ziliaskopoulos, 2001). Other random model components may not clearly translate to real world variability. When trying to compare impact scenarios, it is crucial to understand the amount of deviation created by the randomization processes, otherwise "an analyst has no way to tell whether the change of traffic condition in the compared scenario is strictly due to the scenario or is affected by artifacts introduced by the solution algorithm (Chiu et al., 2010)."

Initial tests for this project began to characterize the randomized variation of link-based statistics. Accounting for the randomness in DTA results can be accomplished through traditional statistical analysis on a link by link basis, aggregated over a subarea, or for the entire network. Attempting to perform a link by link examination for a subnetwork selection process poses a problem.

Once the subnetwork boundaries have been established, the demand information must be extracted from the vehicle trajectories of a full network simulation. Demands from external centroids are summarized to a new boundary centroid by extracting and summing the volumes of links entering the boundary centroid at the edge of the subnetwork in each time period; each vehicle entering the subnetwork also has its destination stored in this demand extraction process. The new, induced time-dependent OD matrix includes only the centroids located within the subnetwork and the newly created boundary centroids.

Subnetwork analysis requires running the network from a cold start with the newly created subnetwork OD demand matrix, which makes comparing the variability of link travel times and volumes nontrivial. The subnetwork assignment produces flows and volumes that contain randomness associated with both the induced boundary demand and the subnetwork simulation. This random variability can be accounted for by treating several simulation results as a sample. This study proposes a methodological framework of statistically comparing OD matrices from multiple scenarios rather than the more common link-level metrics. The error between the base and impact induced subnetwork OD matrices are compared using a two sample equal means test. If the test finds statistical similarity between the errors of the OD matrices, then it may be concluded that the impact scenario does not have significant effects at the boundary of the subnetwork. This indicates that the subarea is sufficiently capable of simulating the impact scenario.

### 3.2 Standard Subnetwork Procedure

The experimental framework is presented here to demonstrate the feasibility of the methodology. The data for this study was received from the Network Modeling Center (NMC) of the Center for Transportation Research (CTR) at the University of Texas at Austin. This analysis uses the Visual Interactive System for Transport Algorithms (VISTA) DTA software. The representative case study used to support this methodology is based on the downtown Austin DTA network extracted from the original Capital Area Metropolitan Planning Organization (CAMPO) five county regional network. The downtown network has been calibrated through an in-depth network characteristic review and validated using available traffic counts. The full network was simulated ten times with base conditions, and ten additional times with each impact scenario. Each of these full network runs was investigated using subnetworks of varying size.

In this study a robust method for characterizing the subnetwork was needed, so the subnetwork elements were selected based on the study performed by Chen et al., 2012 This method of selection, discussed in the literature review in Chapter 2.2, selects network links extending in all directions from the modified link. (The "modified link" is the impact location under study, such as the location of a potential maintenance project, the site of a traffic accident, or a marathon route). A number of additional links, equal to the size parameter, that are topologically connected to the modified link are selected. The larger the size parameter specified, the larger the subnetwork will be. A size parameter of one is one connective link beyond the modified link. Likewise, a size parameter of two is one connective link extending beyond the links included in a size parameter of one. As the size parameter increases by one, the selection includes all contiguous links not included in the previous size parameter selection. Figure 5 shows an example of subnetwork selection using the connected order. Such a selection method is most effective on a homogeneous network structure, but the equal treatment of surrounding links makes it generalizable to common types of networks. The ability to apply this procedure to any network is the goal for tractability of this subnetwork analysis.

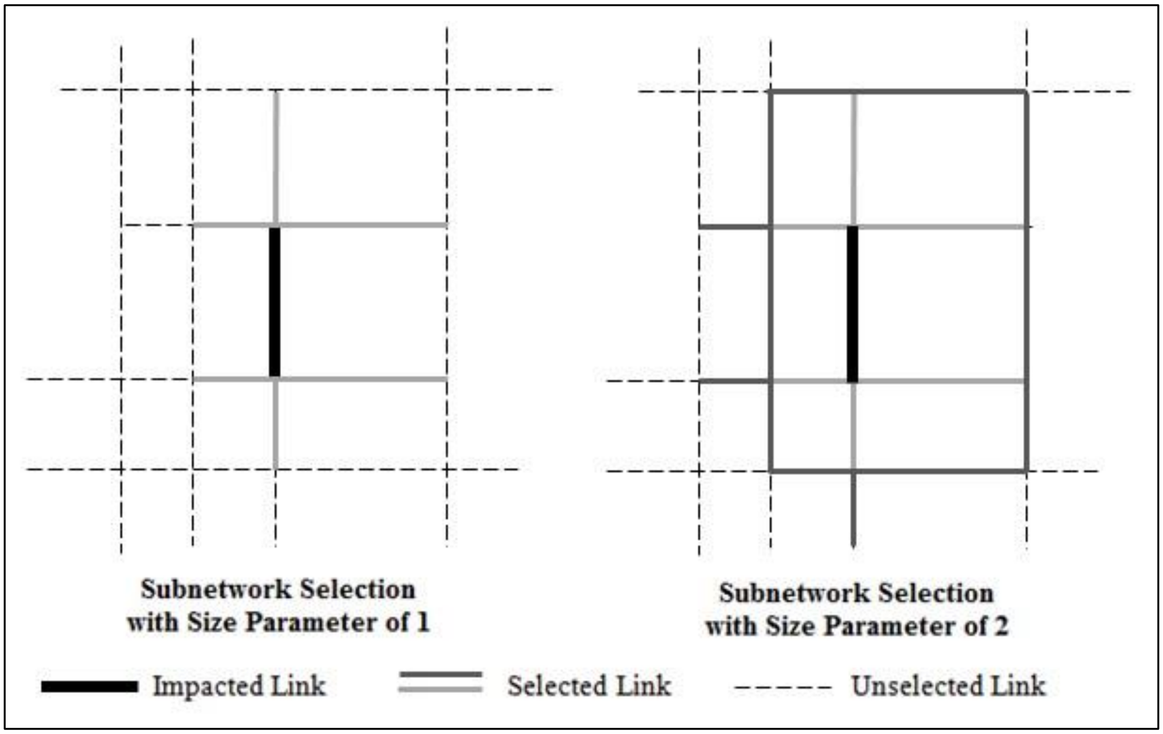



Figure 5: Visualization of the Connected Order Subnetwork Selection Process (from Gemar, 2013)

For the initial subnetwork analysis, this study investigated size parameters of five, seven, and nine because Chen et al., 2012 recommends using size parameters between three and ten to balance the accuracy and efficiency of the subnetwork. The impact scenarios created to assess the limits of the method are defined by three characteristics: roadway identification, link(s) capacity reduction, and the number of modified links. To test the subnetwork size selection, three locations with distinctive characteristics were chosen within the downtown network, including modifications to Guadalupe Street, a southbound one-way arterial with four lanes; 7$^{th}$ Street, an eastbound one-way arterial with four lanes; and 15$^{th}$ Street, an east-west two-way arterial with six lanes (three in each direction). These locations are shown in Figure 6. In addition, capacity reductions of 25 percent, 50 percent, and 100 percent were chosen in combinations involving one link, two links, or three links. These capacity reductions and numbers of links were selected to represent the range within the extremes of potential traffic control scenarios. The same capacity reductions were applied to all links in the scenario. Three options for each of the three impact scenario characteristics led to a total of 81 different scenarios evaluated, each of which was assessed using size parameters of five, seven, and nine. For some of the larger impact scenarios, with three links impacted or 100% capacity reduction, impact scenarios with a size parameter of eleven were investigated because statistical similarity was not found with a subnetwork of connected order nine. These magnitudes of traffic control scenarios were selected to cover typical impact scenarios requiring subarea analysis.

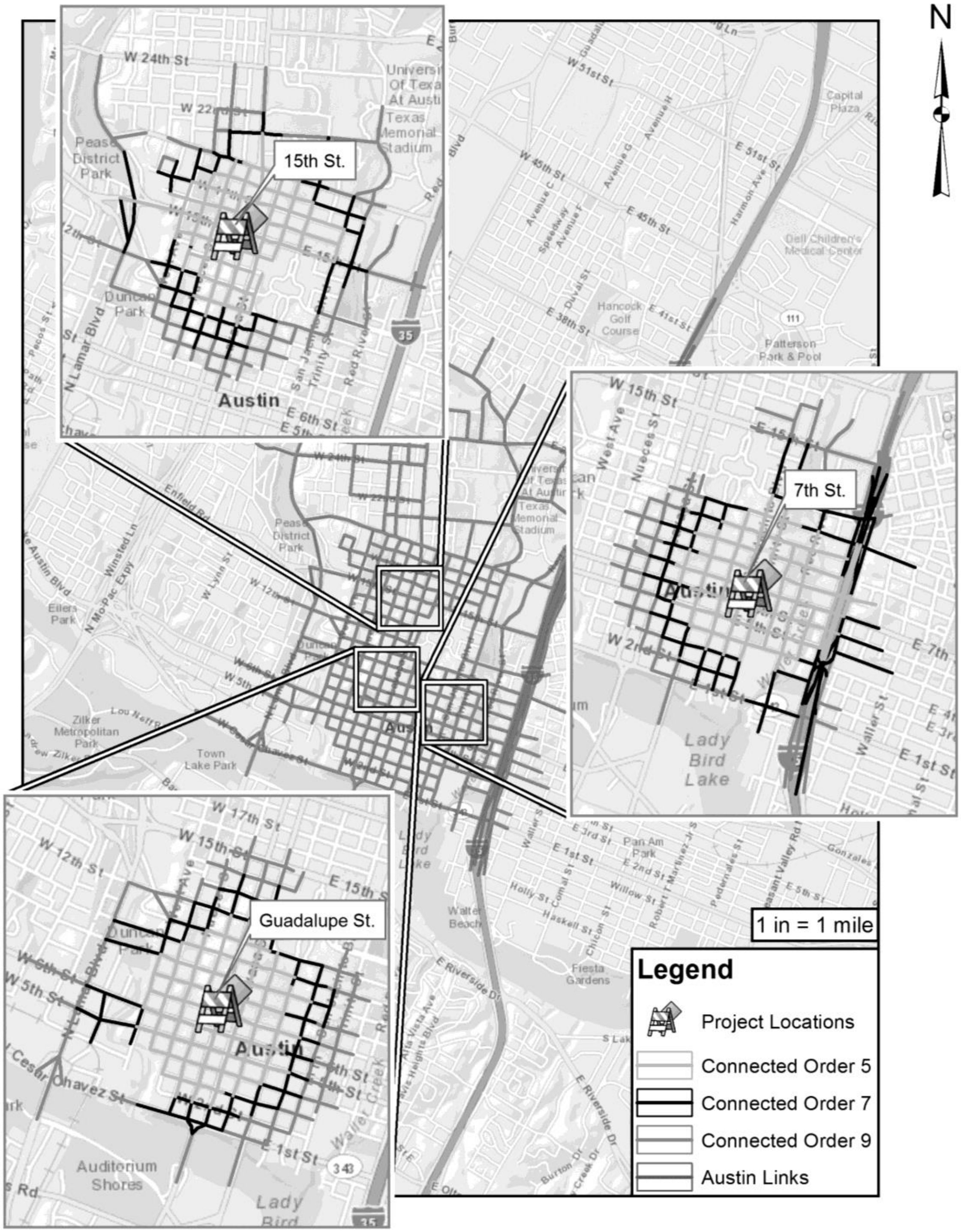



Figure 6: Maps of Impact Scenario Subnetworks in the Downtown Austin Area

The procedure for this experiment included the development of software code for manipulating data and calculating statistics. The procedure that evolved is described in the following eight steps:

***Procedure for Statistical Calculations***

**STEP 1:** Select and extract subnetwork data elements using GIS models

**STEP 2:** Use NMC Java code to extract the induced subarea demand from simulations

**STEP 3:** Extract dynamic OD tables from VISTA and import into spreadsheet

**STEP 4:** Join scenario OD tables with MATLAB code based on time period OD pair

**STEP 5:** Convert joined OD tables to a matrix format as required for SSIM calculations

**STEP 6:** Average values from all the base runs to use as a baseline

**STEP 7:** Calculate error measures comparing each base and impact run to the baseline

**STEP 8:** Compare the base and impact scenario errors using an equal means hypothesis

In this procedure, the term OD table refers to the format that VISTA uses for its demand database. The OD table has three columns that represent the origin, destination, and demand value. An OD matrix is designed so that the rows represent origins, the columns represent destinations, and the cells represent the demand between each row and column OD pair. Only subnetwork OD pairs with demand originating at subnetwork boundary centroids were considered in this analysis because rerouting occurring in the complement network only change the induced demand at the boundary. In other words, subnetwork boundary origin demand values exhibit all of the variability input into the subnetwork between the base and impact scenarios.

As listed in step 4, a procedure was created in MATLAB to match OD pairs with the same time period. This code guarantees that the corresponding demand from each of the different runs of the base scenario and the impact scenario are matched with the identical time-based OD pair for the error calculations. Several demand values from each of the time-dependent OD matrices were not included in all runs (a node was not used during a time period for a particular simulation), which is difficult to account for with typical database join operations. This special code was useful in aligning the appropriate demand values in each subnetwork scenario for accurate error calculations. This was necessary for each of the error measures tested, but extra consideration was needed for the SSIM index.

The MATLAB SSIM index code, written by Wang et al. 2004, requires the standard OD matrix format with rows representing origins and columns representing destinations. The fully joined time-dependent OD table was then parsed and converted to the appropriate format for processing; thus, ensuring the OD matrix dimensions were consistent across base and impact scenario subnetworks. The primary inputs for Wang's code are the subnetwork OD matrices; a few other input parameters for the SSIM code are needed for initialization. The other SSIM factors are described in section 3.3.

The following tests can be completed based on the assumptions that the runs are independent and the simulation output values are identically distributed. The assumption of independence between the runs can be justified by ensuring the individual model runs will be completed from a cold start. Although the network and the demand inputs are the same, the path sets, route assignments, and simulation results will be reset between runs. The assumption of the results coming from the same distribution, in this case normal, can be verified using a Lilliefors or Anderson-Darling test. The Lilliefors test is commonly implemented for small sample sizes (n < 30). The assumption of equal variances between the base and impact scenario sample sets can be verified using a standard F-test. The findings from testing the theoretical assumptions are provided in Chapter 3.5.

After verifying the assumptions, the next step is to determine whether the means of the error measures across the multiple simulations are statistically different between the base and impact scenarios for a given subnetwork. As discussed in Chapter 2.4, these error measures include the root mean square error, the mean absolute percentage error, and the structural similarity index. In order to manage the data processing requirements six representative scenarios were selected to perform ten simulations, and the rest were simulated twice. The six chosen scenarios were Guadalupe Street (one link impact with 25% capacity reduction), Guadalupe Street (two link impact with 50% capacity reduction), Guadalupe Street (three link impact with 100% capacity reduction), 7$^{th}$ Street (two link impact with 50% capacity reduction), 7$^{th}$ Street (three link impact with 100% capacity reduction), and 15$^{th}$ Street (two link impact with 50% capacity reduction). A statistically sufficient comparison requires not only ten runs of the full network base scenario, but also ten runs of each of the full network impact scenarios. A sample size of ten was chosen since each of the models for the full network takes considerable computation time and it has been shown that above this number of samples there are diminishing returns. As a result of the smaller sample size, a two sample t-test for equal means will be used to determine whether the subnetwork boundary demand error is statistically different between the base and impact scenarios. In this test, the null hypothesis is the induced boundary demand error for the base and impact scenario subnetworks are the same.

In addition to the equal means test, a prediction interval was calculated to further examine the hypothesis. For this interval, the base scenario simulations were used to establish a range of boundary demand error that would be expected due to random differences. If the demand extracted at the subnetwork boundary from an impact scenario does not fall within the prediction interval generated from the base scenario, it can be reasonably assumed that this subnetwork is not large enough to capture the influence of the impact scenario. Equation 4 represents the calculation for the prediction interval:

$$PI = \bar{x} \pm t_{\alpha/2,n-1} s \sqrt{1 + \left(\frac{1}{n}\right)} \tag{4}$$

where $\bar{x}$ is the mean, $t_{\alpha/2,n-1}$ is the t-statistic for $\alpha$ level of significance and a sample size of $n$, and $s$ is the sample standard deviation. This concept was compared with the hypothesis test to determine if the prediction interval is robust for evaluating whether the impact scenario demand error measure falls within the expected (model random variation) range of the base scenarios. It is intended that the prediction interval could be used to evaluate subnetwork sizes in place of the more time consuming full hypothesis test. The hypothesis tests require simulating the base and impact scenario multiple times, but the prediction interval method would only require multiple base runs. Since multiple impact scenarios are likely to be tested for traffic control planning process, the prediction interval could save a significant amount of time.

In summary, the three error measures of the subnetwork demand were calculated using the joined OD matrices of each of the ten base and impact scenarios. These error measures were calculated comparing each base and impact scenario to the average of all the base runs. Averaging the demand values in the induced subnetwork OD matrix of all the base runs is the best available estimate for a true baseline. A 95% prediction interval was chosen for the error, and the errors for impact scenarios were compared using a two sample t-test with a 95% confidence level to determine if the difference was statistically significant. The same procedure was carried out for the RMSE, MCAPE and SSIM index.

### 3.3 Structural Similarity Index

The structural similarity index is intended to be a more sophisticated measure of error that captures spatial and temporal effects. Since the SSIM index was originally introduced for image analysis some special considerations must be made for use in an OD matrix analysis context. In particular, selecting the appropriate spatial weights matrices is one of the biggest factors in calculating the SSIM index. For this analysis, the spatial weights matrix represents how strong the correlation is between an OD pair and the surrounding OD pairs in the subnetwork OD matrix. The spatial weights matrix is applied to each OD pair to determine individual SSIM indices, and then the SSIM indices are averaged to determine the overall SSIM index, or MSSIM. This means the spatial weights matrix selected should be general for the entire OD matrix structure. As illustrated in Figure 7, the correlation pattern should emphasize the central feature OD pair of the spatial weights matrix as this is the demand value used to calculate the individual SSIM index, which can be done by giving it the highest weight. In fact, the RMSE and MAPE calculations can be thought of as having a 1x1 spatial weights matrix with a central value of one.

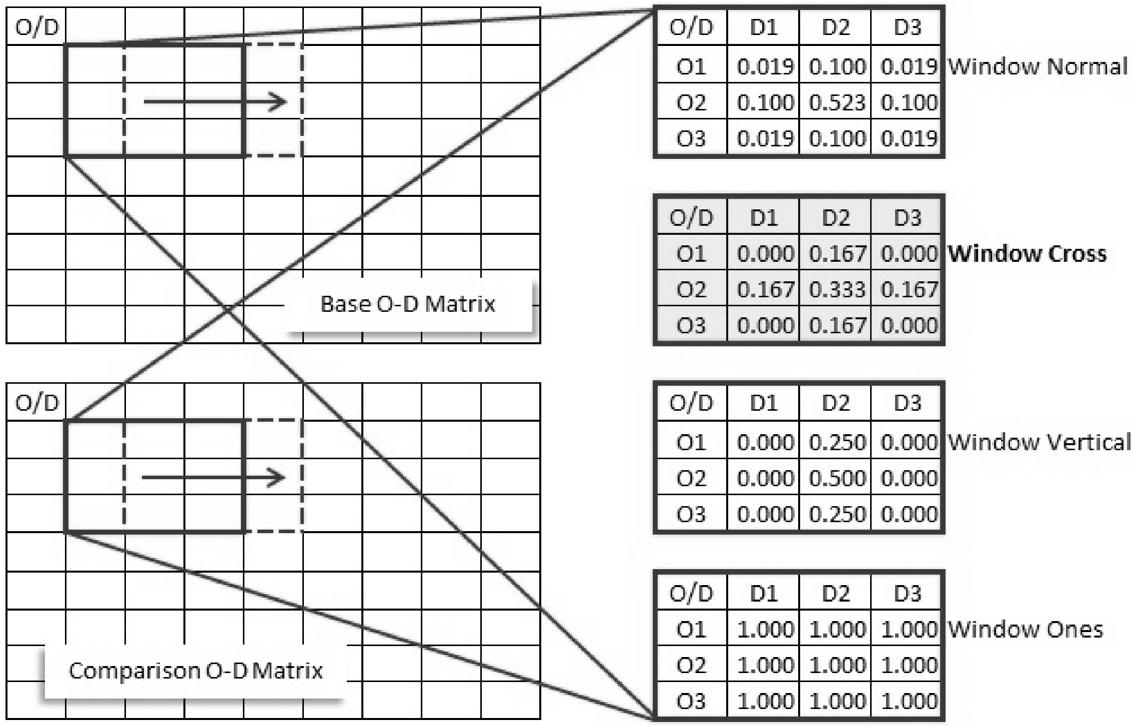
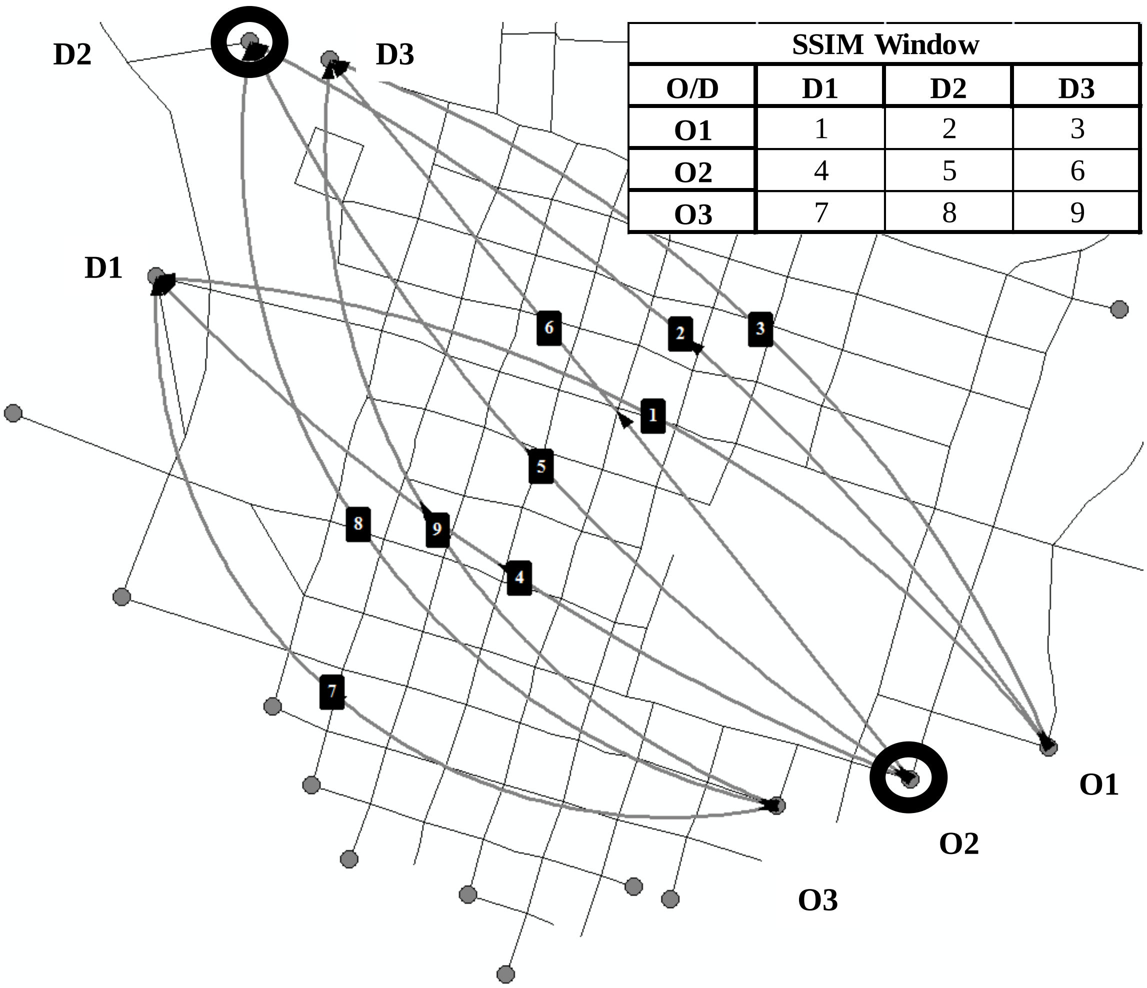



Figure 7: SSIM Spatial Weights Matrices Tested for the Analysis

Figure 7 and Figure 8 together present a representation of the relationship between spatial weight matrices and the corresponding geographic locations of the OD pairs. Potential weights to be tested for a 3x3 matrix are illustrated in Figure 7. A 3x3 window includes only the immediately adjacent origin and destination centroids from the central feature OD pair used to calculate the SSIM index and was chosen because this concept relates to a k-nearest neighbors weighting scheme, where k = 2. K-nearest neighbors is a common GIS concept for a spatial relationship that identifies and selects a specified number, k, of the surrounding GIS features that are the shortest distance from the central feature being analyzed. These features can be points, lines, or polygons, and in this case the features are a centroid (point). The cross window implies that demand *from* the same origin and *to* the same destination should play a dominant role in the analysis.

Alternatively, the vertical window only factors in the demand from nearby origins going to the same destination. The normal window, or a Gaussian distribution matrix, and the "ones" window are two different schemes that include all combinations of surrounding origins and destinations. The central feature OD pair is circled in a bold outline in Figure 8.




| SSIM Window | | | |
|---|---|---|---|
| **O/D** | **D1** | **D2** | **D3** |
| **O1** | 1 | 2 | 3 |
| **O2** | 4 | 5 | 6 |
| **O3** | 7 | 8 | 9 |

Figure 8: Geospatial Representation of Applying a Spatial Weights Matrix to an Origin-Destination Matrix

Figure 8 reveals the importance of the row and column cell location in the OD matrix corresponding to geographic location of the centroid. The figure only contains nodes (points) for boundary centroids for added clarity. O1, O2, O3, D1, D2, and D3 are the representative boundary centroids that depict the location of origins and destinations designated by rows (origins, O) and columns (destinations, D), respectively in Figure 8. The connections 1 through 9 represent the OD pairs that receive the weightings in the 3x3 matrix. Without a proper ordering of the rows and columns in the OD matrix, it is likely the cells surrounding an OD pair would not be geographically related. A unique method of constructing the OD matrix was developed to ensure the geospatial correlation between nearby cells. In order for the windows in Figure 7 to be applied as demonstrated in Figure 8, the subnetwork OD matrix ordering scheme was focused on the subnetwork boundary. Subnetwork boundary centroids were considered first since the application of the error measures was limited to the demand originating at the boundary, where the differences between the base and impact scenario are expected. Subnetwork boundary centroid origins were listed in a clockwise order by traversing the periphery of the subarea starting from the northeastern most point. This order was used for the arrangement of rows in the OD matrix. Then, boundary destinations were listed in the same manner to form the columns of the matrix. The internal destination centroids were investigated to determine if their assigned node identifiers corresponded to geographic location. Generally, sequentially numbered internal destination centroids corresponded with their geographically nearest neighbors (a consideration that should be made as MPOs choose centroid numbers to allow for easier implementation of a variety of geospatial analyses). Internal destinations have far less impact on the SSIM analysis than nearby entry points at the boundary. After addressing the ordering of the subnetwork OD matrix rows and columns, a few more complex considerations can be made for the SSIM analysis.

The basic SSIM index only incorporates spatially related OD cells. In order to incorporate the time dimension into this analysis, another modification to the construction of the origin-destination matrix was considered. Rather than using a matrix organized by geographic location for both origin and destination, only the geographic relationships for origins were used. This again focuses the error measures on capturing the switching of vehicle entry points due to rerouting. The columns, instead, were set up by placing the OD pairs for the previous time period to the left and OD pairs for the subsequent time period to the right of the current time period OD pair. For example, this new OD matrix could have columns with the following order: column 1 from time period 1, column 1 from time period 2, column 1 from time period 3, column 2 from time period 1, column 2 from time period 2, column 2 from time period 2, etc. This is particularly advantageous because of the time-dependent OD matrices used for DTA. Another advantage of only considering the spatial relationship of nearby origins is that it alleviates the issues with ordering the internal destinations in the OD matrix based on their unique identifiers. With this method the SSIM will capture the relationships of vehicles entering at different boundary locations or those boundary locations in different time periods.

With this temporal setup of the SSIM index, the output will have extra columns on either side of the time period of interest. After the SSIM index matrix, or map, is created the SSIM values associated with the current time can be isolated from the previous and subsequent time periods. This creates a SSIM index map with the same number of origins (rows) as boundary origins, and the same number of destinations (columns) as all the destinations within the subnetwork. The spatiotemporal weights matrix used for this test is represented in Figure 9. The MSSIM was calculated for a time only window, space only window, and a time-space window.

| SSIM Temporal Window | | | |
|---|---|---|---|
| **O/D** | **DT1** | **DT2** | **DT3** |
| **O1** | - | - | - |
| **O2** | 1/3 | 1/3 | 1/3 |
| **O3** | - | - | - |
| **SSIM Spatial Window** | | | |
| **O/D** | **DT1** | **DT2** | **DT3** |
| **O1** | - | 1/3 | - |
| **O2** | - | 1/3 | - |
| **O3** | - | 1/3 | - |
| **SSIM Spatiotemporal Window** | | | |
| **O/D** | **DT1** | **DT2** | **DT3** |
| **O1** | - | 1/5 | - |
| **O2** | 1/5 | 1/5 | 1/5 |
| **O3** | - | 1/5 | - |

Figure 9: Comparison of Time and Space Weight Matrices for the Temporal SSIM Index

One last issue with the setup of the origin destination table is the exclusion of the border OD cells. An additional effort was made to include border cells in the SSIM analysis by inserting a fake first row identical to the last row and a fake last row identical to the first row (also done for columns). This enables all of the OD pairs in the matrix to be included in the MSSIM calculation. This also enables a complete loop of the organization of the origins by clockwise orientation around the subnetwork boundary.

Another parameter for the SSIM index is the dynamic range. For an image, the dynamic range represents the value stored in each pixel of an image, which ranges from 0 to 255 for 8-bit color or grayscale. To translate this concept for an OD matrix, the dynamic range should be the maximum of all OD demand values. The dynamic range is used to calculate the constants $C_1$ and $C_2$ to prevent instability in the calculation (i.e. prevent dividing by zero).

Each spatial weights matrix, time period aggregation of simulation results, and even time-based variation will be analyzed to select the optimal SSIM configuration. If the SSIM index has difficulty finding similarity between the base and impact scenarios, an equal distribution of the weights may be used to reduce the error associated with spatial and temporal variability. The added level of sophistication proposed here for the SSIM index has potential application outside this subnetwork methodology. These DTA time-dependent OD matrix and topological considerations contribute to the utilization of this metric for transportation analysis.

### 3.4 Initial Implementation of the Methodology

The statistical equal means tests were implemented for each error measure using one hour and two hour time periods for each of the subnetwork sizes, impact scenarios, and locations. The one hour time period corresponds to the peak hour of the DTA simulation and the two hour time period corresponds to the peak period. In these cases, the induced boundary demands were summed over the peak hour and peak periods. Considerations were made for these time periods to avoid the initial loading of the DTA network and the unloading during the end of the simulation, as these output are less reliable. This analysis revealed that the RMSE was the best tool for determining subnetwork size because it was able to identify a transition to statistical similarity of the induced boundary demand as the subnetwork size increased. The SSIM index found the base and impact subnetwork OD matrices to be statistically different no matter what size parameter, spatial weights, or time period was used. This indicated that the SSIM index was too sensitive to the error between base and impact scenarios. The MAPE was also found to be too sensitive to the error and not useful for determining a subnetwork size. Since previous studies have shown the chosen ranges of subnetwork sizes to be sufficient for capturing network alterations, the MAPE and SSIM index were eliminated from the study.

The RMSE gave the most promising results for comparing induced demand error associated with incrementally increasing the size of the subnetwork. For each of the scenarios, the RMSE indicates that boundary demands for smaller sized subnetworks are statistically different compared to the base, and for larger subnetworks, these demands are statistically the same. This transition to statistical similarity indicates that the subnetwork is large enough to contain a majority of the traffic impacts. The RMSE results demonstrate an ability to identify this threshold; therefore, it is a viable candidate for a subnetwork size metric.

Using RMSE as a base for comparison the three different weighting concepts in Figure 9 were tested. To address the oversensitivity of the spatial only SSIM index, the subnetwork impact scenario with the smallest recommended size from the RMSE results was used. The Guadalupe impact scenario with a capacity reduction of 25% on one link found similarity between the base and impact scenarios at a subnetwork of size parameter five. It was hypothesized that incorporating more time periods in the SSIM analysis would increase the sensitivity of the time periods used for simulation results. Therefore, the proposed temporal SSIM was tested on a number of time period subsets to determine if any time period would enhance the tractability of the SSIM index. The time periods referenced here are 15 minute intervals of the simulation period. For instance, time period 2 through 5 represent the standard peak hour for this study, or 30 minutes into simulation through 90 minutes into simulation. Table 1 documents the results of the spatiotemporal SSIM from the different time periods.

Table 1: Statistical Tests of the Proposed Spatiotemporal SSIM Index for 10 Base Scenarios and 10 Impact Scenarios across Different Time Periods

| | | | | | Hypothesis Testing | | | |
|---|---|---|---|---|---|---|---|---|
| Scenario | Location | Subnetwork Size (Order) | Impact Size | Time Periods | Equal Variance* | Normality Lilliefors** | Normality A-D** | Equal Mean*** |
| Base | Guadalupe St | 5 | 1, 25% | All | Y | Y | Y | Y |
| Impact | Guadalupe St | 5 | 1, 25% | All | | Y | Y | |
| Base | Guadalupe St | 7 | 1, 25% | All | Y | Y | Y | Y |
| Impact | Guadalupe St | 7 | 1, 25% | All | | Y | Y | |
| Base | Guadalupe St | 9 | 1, 25% | All | Y | Y | Y | Y |
| Impact | Guadalupe St | 9 | 1, 25% | All | | Y | Y | |
| Base | Guadalupe St | 5 | 1, 25% | 2 to 5 | Y | Y | Y | N |
| Impact | Guadalupe St | 5 | 1, 25% | 2 to 5 | | Y | Y | |
| Base | Guadalupe St | 7 | 1, 25% | 2 to 5 | Y | Y | Y | N |
| Impact | Guadalupe St | 7 | 1, 25% | 2 to 5 | | Y | Y | |
| Base | Guadalupe St | 9 | 1, 25% | 2 to 5 | Y | Y | Y | N |
| Impact | Guadalupe St | 9 | 1, 25% | 2 to 5 | | Y | N | |
| Base | Guadalupe St | 5 | 1, 25% | 4 to 7 | Y | Y | Y | Y |
| Impact | Guadalupe St | 5 | 1, 25% | 4 to 7 | | Y | Y | |
| Base | Guadalupe St | 7 | 1, 25% | 4 to 7 | Y | Y | Y | N |
| Impact | Guadalupe St | 7 | 1, 25% | 4 to 7 | | Y | Y | |
| Base | Guadalupe St | 9 | 1, 25% | 4 to 7 | Y | Y | Y | N |
| Impact | Guadalupe St | 9 | 1, 25% | 4 to 7 | | Y | Y | |
| Base | Guadalupe St | 5 | 1, 25% | 6 to 9 | Y | Y | Y | N |
| Impact | Guadalupe St | 5 | 1, 25% | 6 to 9 | | Y | Y | |
| Base | Guadalupe St | 7 | 1, 25% | 6 to 9 | Y | Y | Y | Y |
| Impact | Guadalupe St | 7 | 1, 25% | 6 to 9 | | Y | Y | |
| Base | Guadalupe St | 9 | 1, 25% | 6 to 9 | Y | Y | Y | N |
| Impact | Guadalupe St | 9 | 1, 25% | 6 to 9 | | Y | Y | |
| Base | Guadalupe St | 5 | 1, 25% | 3 to 5 | Y | N | N | N |
| Impact | Guadalupe St | 5 | 1, 25% | 3 to 5 | | Y | Y | |
| Base | Guadalupe St | 7 | 1, 25% | 3 to 5 | Y | Y | Y | N |
| Impact | Guadalupe St | 7 | 1, 25% | 3 to 5 | | Y | Y | |
| Base | Guadalupe St | 9 | 1, 25% | 3 to 5 | Y | Y | Y | Y |
| Impact | Guadalupe St | 9 | 1, 25% | 3 to 5 | | Y | N | |
| Base | Guadalupe St | 5 | 1, 25% | 2 to 9 | Y | Y | Y | N |
| Impact | Guadalupe St | 5 | 1, 25% | 2 to 9 | | Y | N | |
| Base | Guadalupe St | 7 | 1, 25% | 2 to 9 | Y | N | Y | N |
| Impact | Guadalupe St | 7 | 1, 25% | 2 to 9 | | Y | Y | |
| Base | Guadalupe St | 9 | 1, 25% | 2 to 9 | N | Y | Y | N |
| Impact | Guadalupe St | 9 | 1, 25% | 2 to 9 | | Y | Y | |

* Y = Accept $H_o$: $\sigma_1^2 = \sigma_2^2$; N = Reject $H_o$, conclude $H_a$: $\sigma_1^2 \neq \sigma_2^2$
** Y = Accept $H_o$: Distribution is normal; N = Reject $H_o$, conclude $H_a$: Distribution is not normal
*** Y = Accept $H_o$: $\mu_1^2 = \mu_2^2$; N = Reject $H_o$ conclude $H_a$: $\mu_1^2 \neq \mu_2^2$

Unfortunately, the results in Table 2 indicate that the temporal SSIM index performs inconsistently. For the peak hour, which was also used for the RMSE analysis, all subnetwork sizes were found to be dissimilar. In other time periods, the statistical test performed on the temporal SSIM index indicates that the incremental increase in subnetwork size does not capture more of the rerouting at the boundary. In particular, time periods 4 to 7 (or an hour into the simulation to two hours into the simulation) found a subnetwork of size 5 to be adequate, but larger subnetworks to not be sufficient. Also, time periods 6 to 9, the second hour of the peak period, found size 7 to be adequate, but size parameters 5 and 9 were not. This sort of erratic behavior is not acceptable for the type of metric necessary for these predictions. It was hypothesized that if the temporal SSIM predicted results similar to RMSE, and possibly even identify smaller, adequate size parameters, that it would have been a more robust metric. Adding the time domain, which is pertinent to DTA analysis, would have allowed the SSIM index to incorporate the temporal effects of DTA simulation. However, the results from the statistical procedure find the spatiotemporal SSIM to be as unreliable as the MAPE and ordinary SSIM index. Therefore, it was also be removed from further testing.

These statistics that were abandoned may still have a place in transportation analysis. However, their potential for addressing subnetwork boundary rerouting has limited their usefulness for this methodology. The SSIM index had the potential to provide another component to the analysis by incorporating spatial effects, which traditional statistical methodologies do not address despite their prevalence in transportation data. This investigation did find that a spatial weights matrix with a cross-shaped distribution incorporated the appropriate spatial relationships, and should be implemented if used for further OD matrix study. The MAPE error measure had the potential to normalize the changes in boundary demand, which may be beneficial for OD matrices associated with heterogeneous network characteristics. This investigation found that it may be useful to censor the MAPE by taking the minimum of the individual MAPE values and 100%. The

MCAPE can prevent overestimation of the OD matrix error. Further development of the subnetwork methodology will focus on the use of the RMSE.

### 3.5 Refining the Statistical Subnetwork Characterization

The preliminary investigation of the error measures, evaluated a subset of scenarios representative of a comprehensive range of possible network modifications. The RMSE was found to produce the most valuable results during the one hour period of the simulation time for the trial scenarios, representing the peak hour. Practitioners have also suspected that this time period is the most reliable portion of the simulation output. The trial scenarios included a 25 percent capacity reduction to one link, a 50 percent reduction to two consecutive links, and a 100 percent capacity reduction across three consecutive links. To control for location, all three scenarios were tested along Guadalupe Street. To control for impact scope, the mid-range scenario characteristic of a 50 percent capacity reduction to two links was tested at all three locations. A second, 100 percent capacity reduction was also tested to further investigate the most substantial impact scenario.

Table 2: Statistical Tests of the RMSE for 10 Base and 10 Impact Scenarios during the Peak-Hour

| Scenario | Location | Subnetwork Size (Order) | Impact Size | Capacity Reduction | Hypothesis Testing | | | | Prediction Interval | | | Impact Runs Within Base Range |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Equal Variance* | Normality Lilliefors** | Normality A-D** | Equal Mean*** | Lower | Upper | Range | |
| Base | Guadalupe St | 5 | 1 | 25 | Y | Y | Y | Y | 1.73 | 4.81 | 3.08 | 10 |
| Impact | Guadalupe St | 5 | 1 | 25 | | Y | Y | | 2.42 | 4.27 | 1.85 | 9 |
| Base | Guadalupe St | 7 | 1 | 25 | Y | Y | Y | Y | 1.57 | 4.37 | 2.80 | 10 |
| Impact | Guadalupe St | 7 | 1 | 25 | | Y | Y | | 2.33 | 3.90 | 1.57 | 9 |
| Base | Guadalupe St | 9 | 1 | 25 | Y | Y | Y | Y | 1.64 | 3.40 | 1.76 | 10 |
| Impact | Guadalupe St | 9 | 1 | 25 | | Y | Y | | 1.84 | 3.33 | 1.49 | 10 |
| Base | 15th St | 5 | 2 | 50 | Y | Y | Y | N | 1.89 | 3.60 | 1.72 | 10 |
| Impact | 15th St | 5 | 2 | 50 | | Y | Y | | 2.19 | 4.05 | 1.86 | 10 |
| Base | 15th St | 7 | 2 | 50 | N | N | N | Y | 0.33 | 5.30 | 4.97 | 10 |
| Impact | 15th St | 7 | 2 | 50 | | Y | Y | | 2.45 | 3.75 | 1.30 | 5 |
| Base | 15th St | 9 | 2 | 50 | Y | Y | Y | Y | 1.24 | 3.15 | 1.91 | 9 |
| Impact | 15th St | 9 | 2 | 50 | | Y | Y | | 1.71 | 3.27 | 1.56 | 10 |
| Base | 7th St | 5 | 2 | 50 | Y | Y | N | N | 1.93 | 2.97 | 1.04 | 5 |
| Impact | 7th St | 5 | 2 | 50 | | Y | Y | | 2.31 | 3.52 | 1.21 | 8 |
| Base | 7th St | 7 | 2 | 50 | Y | Y | Y | N | 1.36 | 2.32 | 0.96 | 6 |
| Impact | 7th St | 7 | 2 | 50 | | Y | Y | | 1.58 | 2.74 | 1.16 | 9 |
| Base | 7th St | 9 | 2 | 50 | N | Y | Y | Y | 0.85 | 2.16 | 1.31 | 10 |
| Impact | 7th St | 9 | 2 | 50 | | Y | Y | | 1.15 | 1.77 | 0.62 | 6 |
| Base | Guadalupe St | 5 | 2 | 50 | N | Y | N | Y | 1.53 | 4.69 | 3.16 | 10 |
| Impact | Guadalupe St | 5 | 2 | 50 | | Y | Y | | 2.40 | 3.90 | 1.50 | 8 |
| Base | Guadalupe St | 7 | 2 | 50 | Y | Y | Y | Y | 1.53 | 4.51 | 2.98 | 10 |
| Impact | Guadalupe St | 7 | 2 | 50 | | Y | Y | | 2.23 | 4.16 | 1.94 | 9 |
| Base | Guadalupe St | 9 | 2 | 50 | N | N | Y | Y | 1.57 | 3.90 | 2.34 | 10 |
| Impact | Guadalupe St | 9 | 2 | 50 | | Y | Y | | 2.20 | 3.25 | 1.05 | 6 |
| Base | 7th St | 5 | 3 | 100 | Y | N | N | N | 1.92 | 2.93 | 1.01 | 0 |
| Impact | 7th St | 5 | 3 | 100 | | Y | Y | | 8.98 | 10.68 | 1.70 | 0 |
| Base | 7th St | 7 | 3 | 100 | Y | N | Y | N | 1.23 | 2.81 | 1.58 | 6 |
| Impact | 7th St | 7 | 3 | 100 | | Y | Y | | 2.12 | 3.36 | 1.24 | 4 |
| Base | 7th St | 9 | 3 | 100 | Y | Y | Y | N | 0.84 | 2.18 | 1.34 | 10 |
| Impact | 7th St | 9 | 3 | 100 | | Y | Y | | 1.36 | 2.17 | 0.81 | 6 |
| Base | Guadalupe St | 5 | 3 | 100 | Y | Y | Y | N | 1.55 | 4.50 | 2.95 | 0 |
| Impact | Guadalupe St | 5 | 3 | 100 | | Y | Y | | 6.01 | 7.56 | 1.56 | 0 |
| Base | Guadalupe St | 7 | 3 | 100 | Y | Y | Y | N | 2.23 | 4.78 | 2.56 | 7 |
| Impact | Guadalupe St | 7 | 3 | 100 | | Y | Y | | 3.43 | 5.64 | 2.21 | 5 |
| Base | Guadalupe St | 9 | 3 | 100 | Y | Y | Y | Y | 1.25 | 4.00 | 2.75 | 10 |
| Impact | Guadalupe St | 9 | 3 | 100 | | Y | Y | | 2.16 | 3.70 | 1.54 | 7 |

* Y = Accept $H_o$: $\sigma_1^2 = \sigma_2^2$; N = Reject $H_o$, conclude $H_a$: $\sigma_1^2 \neq \sigma_2^2$
** Y = Accept $H_o$: Distribution is normal; N = Reject $H_o$, conclude $H_a$: Distribution is not normal
*** Y = Accept $H_o$: $\mu_1 = \mu_2$; N = Reject $H_o$ conclude $H_a$: $\mu_1 \neq \mu_2$

Table 2 summarizes the RMSE statistical test results for the peak hour of all 10 simulations of the base scenario and 10 simulations of the impact scenario. The table can be interpreted using the first five columns to establish the scenario: base or impact; location by street name; subnetwork selection size parameter, or connected order; impact size in number of links; and impact magnitude in percent capacity reduction. As seen in Table 2 the majority of scenarios passed the equal variance, Lilliefors, and Anderson-Darling tests, which indicates that the independent and identical distribution assumption is correct. The last four columns represent the results of the prediction interval analysis. The lower bound, upper bound, and range of the prediction interval of the base scenario induced subnetwork demand errors are listed first; the last column represents the number of impact run errors that fell in the range of the base simulation error prediction interval. A greater number of impact runs within the base range is intended to indicate that the subnetwork may be sufficient. The prediction interval, although valuable in concept, has a limited application for making subnetwork recommendations because the results were not as consistent or meaningful as the equal means hypothesis test.

The most important column in Table 2 is the equal means test results. The RMSE indicated that for minor impact scenarios (1 link, 25% capacity reduction) the base and impact subnetwork demand errors were equivalent at all subnetwork sizes, indicating that a subnetwork of size parameter 5 is sufficient. For larger impacts, the transition from statistical similarity to a statistical difference in boundary demand error was also captured. For instance, the 7$^{th}$ Street, 2 link, 50% capacity reduction scenario indicated that a size parameter of 7 was insufficient, but 9 was sufficient. It was anticipated that the threshold size for each scenario could be established using these comparisons, providing a means for determining the appropriate subnetwork selection necessary for capturing the majority of traffic impacts resulting from a network modification. The results in Table 2 suggesting an incremental change in the sufficiency of a subnetwork as the size grows, corroborate intuition and imply that the proposed statistical analysis can account for and eliminate the random variation associated with the model.

Once the results had shown that the RMSE is capable of identifying when there is no longer a statistically significant difference between the base and impact subnetwork boundary demand, the RMSE was used to evaluate all scenario combinations to derive systematic recommendations. This includes the scenarios that were only simulated twice. Although, two simulations does not provide a detailed distribution, two runs do protect against a single value coming from an outlier in the distribution and alleviated the burden of computational time needed for more simulations. The results from this extended analysis are shown in Figure 10. Figure 10 can be understood by reading the number of links modified for a given scenario (to the left) and the percent capacity reduction (on top), and tracing each to the point of intersection within the region of the recommended subnetwork size parameter. For example, a 2 link impact with 50% capacity reduction corresponds to a size parameter of 7, which corresponds to the results from the trial scenarios in Table 2. These recommendations may be used by practitioners to construct subareas for analysis corresponding to real world scenarios. However, the limitation of the connected order is that it does not address the peripheral structure of the subnetwork. Manual editing of the boundary should be used to ensure that connectivity of corridors is maintained around the edge of the subnetwork as well as not disconnecting approaches of intersections.

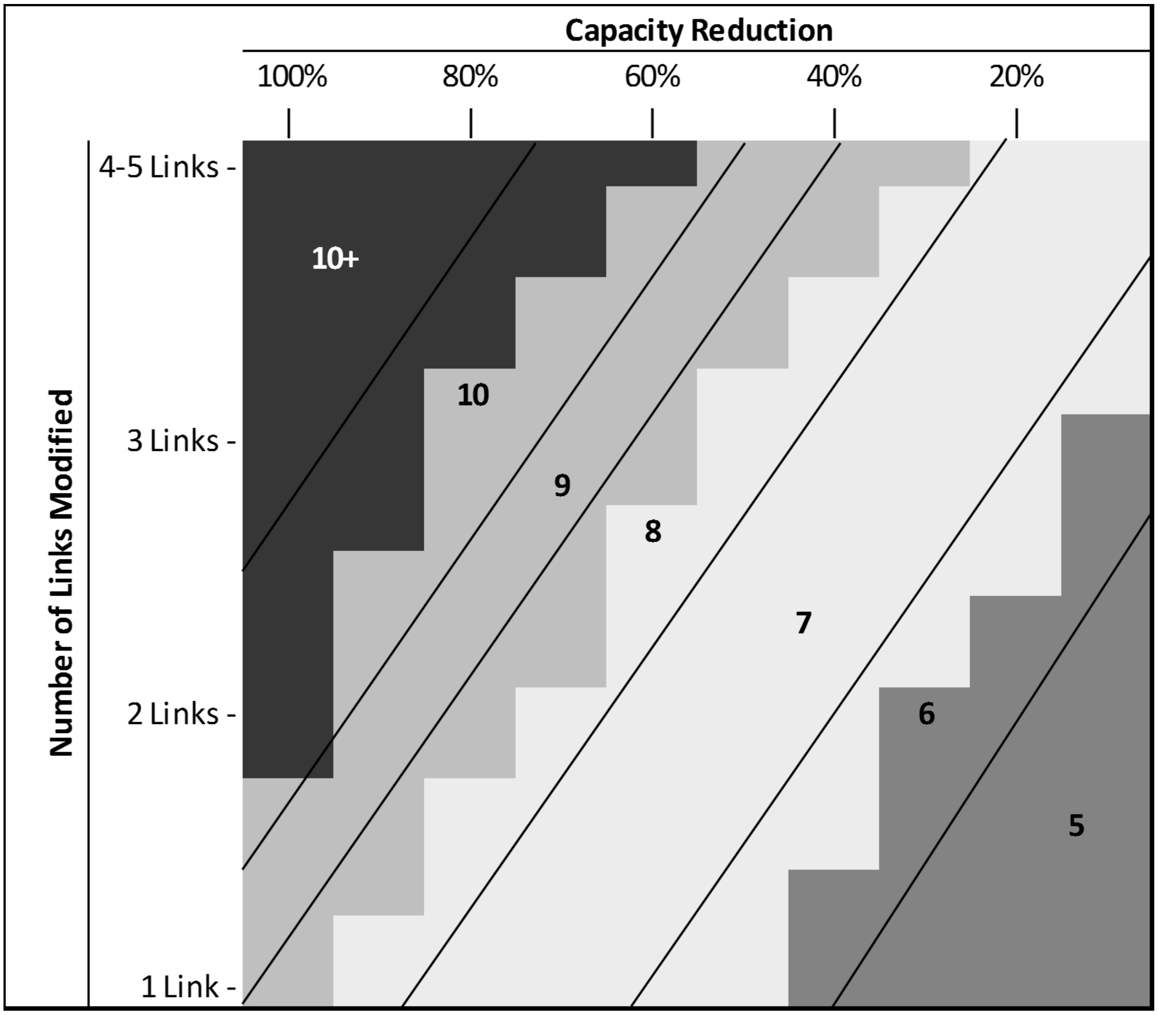



Figure 10: Detailed Results of the Subnetwork Size Evaluation (from Gemar, 2014)

Through an evaluation of both traditional and unconventional statistical measures, the root mean square error was found to be the most useful measure for evaluating the subnetwork size relative to an impact scenario. Subsequently, a set of guidelines and recommendations has been established to determine the approximate subnetwork size required to evaluate a network modification based on the designated number of links to be altered and the associated capacity reduction. Final recommendations produced from the RMSE investigation of all representative scenarios are presented in Figure 10.

In order for any statistic to provide insight into this analysis for required size of subnetwork, it must be able to reveal changes through a hypothesis test comparing base and impact scenario demand errors. That is, it should show that boundary demands for insufficiently-sized subnetworks are statistically different, and for an adequate subnetwork size these demands are statistically the same. This transition to statistical similarity indicates that the subnetwork is large enough to contain a majority of the local traffic impacts, and it is assumed that the subnetwork size is sufficient. By capturing this transition rather than comparing outputs directly, the random variation associated with stochastic model components can be separated from the analysis. Removing this variation allows the user to identify potential extent of impacts from imposed network modifications.

### 3.6 SUMMARY

This chapter covers the evolution of the methodology for using subnetwork OD comparison as a means to identify an appropriate subnetwork size. The random variation of model results required a statistical test for comparing OD errors rather than a measure of the absolute difference. DTA model complexity lends itself to this benchmarking methodology based on simulations rather than an optimization framework. The procedure for the statistical test using each error measure (RMSE, MAPE, and SSIM) is documented in Chapter 3.2. The complexity of the SSIM index required further explanation of the geographic and time based methods used to capture the OD error. A brief summary of the preliminary sensitivity analysis results are used to demonstrate what makes an error measure capable of identifying the appropriate subnetwork size. Then, further scenarios were used to provide a solid foundation for subnetwork size recommendations based on the RMSE metric. The results of this analysis are summarized in Figure 10. Since the RMSE found an incremental decrease in error between the base and impact scenarios as the size of the subarea increased, this error measure will be used as the basis for a potential statistical prediction model. An overview of the development of this methodology is presented in Figure 11.

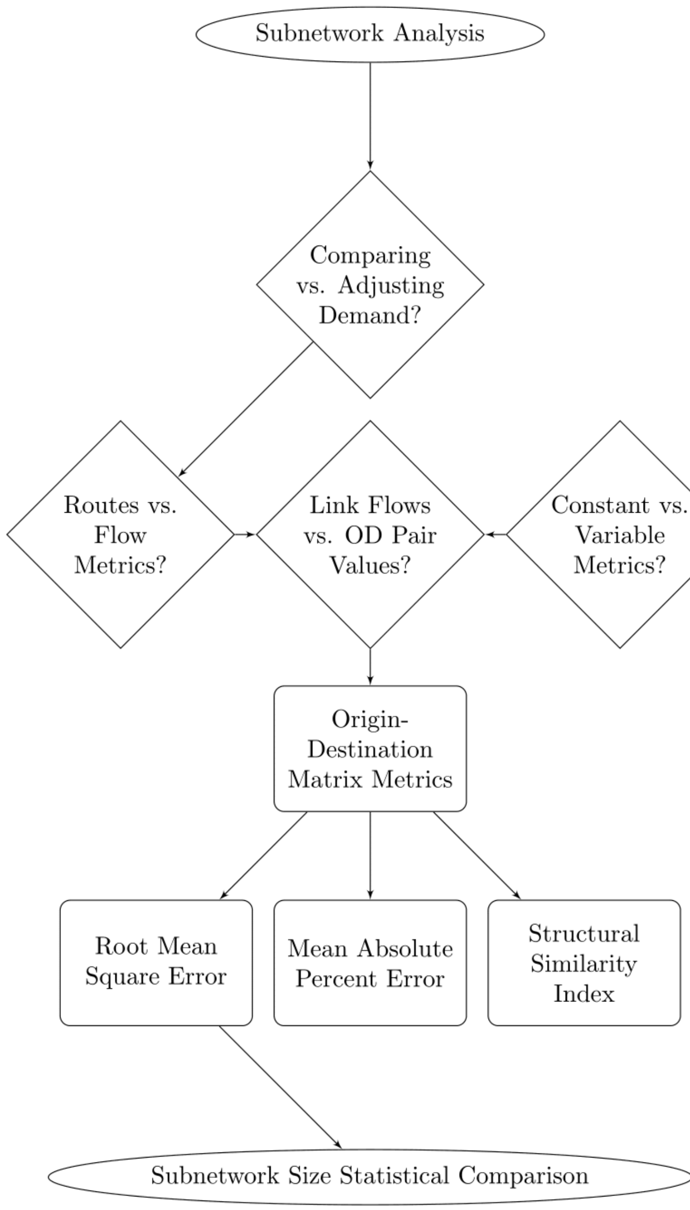



Figure 11: Flow Chart for the Subnetwork Comparison Methodology

# Chapter 4: Prediction

Statistical comparison of the subarea origin-destination (OD) demand matrices proved to be an appropriate method for finding a sufficient subnetwork. By examining flows in the base and the impact network scenarios the root mean square error metric of rerouting can be used to identify statistical similarity of the subnetwork inputs. For a particular network configuration, the same statistical comparison could be applied to derive recommendations similar to the results from Figure 10. However, the effort required to perform this analysis may be prohibitive and it is hypothesized that certain characteristics of the impact scenario and the network may be used to predict this error without multiple network simulations. Specifying this relationship may help to characterize an appropriate subarea in significantly less time and effort. This technique may also have the potential to adapt this subnetwork characterization methodology to a variety of network types.

The comparison technique is a trial and error process for analyzing subnetwork OD matrices that can support subnetwork recommendations based on average results for a given scenario. It may be possible to produce a prediction of error expected at the subnetwork boundary due to vehicles rerouting. The statistical tests used to investigate the comparison method indicate that the data also fulfill the assumptions for a linear regression model. Since the trial and error approach of the comparison technique requires multiple simulations of both the base and impact scenarios, a large portion of the computational effort could be eliminated through the use of a predictive model. A closed form prediction of error associated with subnetwork size is desirable to eliminate the effort required by the trial and error process. The predictive model should relate multiple impact scenario variables to the induced boundary demand error. This technique has the potential to be a more robust method for designating a subnetwork size. This chapter will cover the methodology for building a predictive model and how the model may be used to establish an appropriate subnetwork size based on an error threshold.

### 4.1 Transportation Forecast Models

Historically, results from traffic simulation models have been a primary means for developing traffic impact analyses. The process is as follows: a simulation network is constructed, the before and after case scenarios are simulated multiple times, and averages of the results are compared to identify changes in performance metrics. These performance metrics are commonly fundamental traffic flow characteristics including density, speed, and delay. A more intuitive interpretation is generated using the level of service concept that relates these principle metrics to a letter grade scale. It is a novel idea to attempt to determine the relationship between the imposed network impacts and the changes experienced in the simulation results. In addition, performing such an analysis to examine the spatial scale at which significant changes occur introduces a more sophisticated form of traffic analysis. Now questions can be asked regarding what spatial extent is necessary for the traffic simulation to be an effective estimate of the changes to a network.

Identifying the appropriate spatial scale that captures the impacts of a given scenario opens up more possibilities for enhancing the accuracy of traffic simulation. This methodology provides a theoretical framework for multi-resolution modeling. Statistical analysis, like the concepts used to estimate level of service from multiple runs, may also be used to address when the different levels of simulation detail– macroscopic, mesoscopic, and microscopic – should be applied. That is, these methods can predict where aggregate measures of rerouting, more detailed rerouting, and second by second vehicle trajectories are necessary. Attempting to use a statistical hypothesis test to answer this question, recognizes the significant role played by real world variability. Taking this methodology one step further with a regression model designed to predict error associated with alternative subnetwork sizes provides insight to the variance of network measure outputs with the input variables.

This direction for the methodology was also guided by analyzing the spatial scale associated with different network types. The comparison methodology was verified with a fairly homogeneous, well-connected grid network. While this type of network is common for most urban areas, this subnetwork procedure would be valuable if it could also be used for rural networks and mature freeway network applications. Mature freeway networks, such as multiple ring roads, common to many of the largest cities, and rural networks represent the two extremes of network connectivity. In order to address these different types of connectivity, alternative techniques of subnetwork selection were hypothesized. For instance, on freeway networks one could conceive of a connected order related to intersections of freeways, rather than treating freeways as being similar to collector or local links; or using the connected order selection method for links with capacity greater than 2,000 first, then applying a second connected order selection for all links. These techniques would have led to a similar comparison analysis and associated recommendations, but would not have added to the methodology. Instead, developing a new research method could allow for a greater understanding of how the network type influences simulation behavior. The potential for using a characterization of the network topology as a variable can generalize this process for recommending a sufficient subnetwork size.

Previously addressed considerations for the application of linear regression to transportation will help clarify some of the concepts used for this analysis. The basic objective of linear regression is to estimate the value of a dependent variable using known independent, or predictor, variables. In this case, rather than using multiple observations of the root mean square error (RMSE) of the subnetwork OD matrix to provide a sufficient subnetwork size, one can use known independent variable values and a regression equation to create an estimate of the RMSE. These independent variables would ideally be characteristics of the network and impact scenario that influence the propagation of rerouting. There are four major assumptions that are required for using a linear regression model. 1) The dependent variable should have a linear relationship with the independent variable. 2) The error of the predicted value should be normally distributed. 3) The data should be independent. 4) The variance of the predictor should be constant across the range of the independent variable. Each of these assumptions must be verified for the dataset of subnetwork OD matrix errors.

There has been widespread adaptation of linear regression techniques in transportation planning and engineering. The standard four step transportation planning model often uses linear regression models to predict numbers of trips beginning and ending in traffic analysis zones (Martin and McGuckin, 1998). More recent advances have adapted linear regression techniques to estimate time-dependent OD trip matrices based on known traffic counts (Bierlaire and Critten, 2004; Zhou et al. 2003). These studies related the deviation in actual and historical OD flow to sensor data of traffic volumes. This application provides further support that the deviations in OD matrices are normally distributed and may be predicted using other known variables. Given the common use of linear regression for transportation analysis, it appears suitable for the analysis of subnetwork OD matrices.

The motivation behind these countless applications of linear regression in transportation is the capability of predictive models to forecast future conditions. Large data collection efforts made available via mobile telecommunication systems are getting the industry closer to an accurate depiction of the state of the transportation system. However, most decisions transportation planners and engineers face require analysis of the state of the transportation systems years into the future. Current model calibration data are usually expensive and the overall model value is determined by the ease and cost of predicting future predictor variable values. Time series analysis tools are often used to develop predictive models for parameters whose values change with time, but many of these tools rely on ordinary least squares estimation. Beyond time series analysis, linear regression allows for enhanced predictive power by exploring relationships between the predicted variable and almost any predictor variable hypothesized by the analyst. For the particular problem described in this dissertation, the predictive regression approach means that it can be a more robust method for creating a subnetwork. The variables included in the recommended model will make the procedure more applicable to different types of networks and scenarios.

### 4.2 Constructing a Basic Model

The first consideration for creating a linear model is the potential variables that intuitively influence the dependent variable. The RMSE of the subnetwork OD matrices is the dependent variable, or response variable. The model will assume the RMSE is linearly influenced by the proposed factors related to the impact scenario. This makes sense because the error expected at each OD pair is a result of vehicles changing paths due to changes inside the subnetwork, and the amount of path changing should be proportional to the scale of the changes or impact. Therefore, the most valuable predictor, or explanatory, variables would be characteristics of the impact scenario and the subnetwork. The initial three predictor variables that were considered are the capacity reduction on the modified links, the number of links modified, and the subnetwork size parameter. Further tests of other variables determined their significance, if any, for predicting the RMSE.

The data used for calibration of this model was generated during the comparison test procedure. Using the same data allows for a comparison of the previous recommendations and the capability of the linear model. Again, the predicted error is the difference between the demands recorded at the subnetwork boundary from runs of the entire network and demands recorded at the subnetwork boundary from runs of the entire network with the impact scenario. A RMSE of one indicates that each boundary centroid was off by an average demand of one vehicle in each time period. This error is caused by the rerouting of vehicles to avoid the capacity reduction in favor of a shorter path. In total there were 306 distinct RMSE calculations associated with a variety of impact scenario characteristics. This sample data will be used to build a predictive model for the overall magnitude of rerouting at the subnetwork boundary.

The creation of a subnetwork requires a method to determine the area around the network modification that experiences significant change. While changes may occur throughout the entire network, the comparison of subnetwork OD matrices has provided evidence that these changes decrease with distance from the network modification. This is an important step in understanding the spatial scale of network effects and making recommendations for appropriate subnetwork sizes. However, although potentially robust, a trial and error approach would rely on average values based on hundreds of runs of a full network. This procedure would be time consuming and would not be feasible each time a new network is used and a new user wants to calibrate to their needs. Based on the statistical tests used for the comparison method, the next step in the statistical analysis is to build a predictive model that may be used to estimate the relationship between subnetwork size and error. Preliminary testing of the comparison model provides a statistical foundation for the predictive linear regression model.

In order to use an equal means t-test on the error measures between the base and impact scenarios, the error had to pass a normality test. The majority subnetwork boundary demand error for the sample simulations in both the base and impact scenarios were found to be normally distributed. The limited number of runs used for the data analysis may be the cause of the few cases where the error did not appear to be normally distributed. Further runs of the same scenario were avoided due to computational time required for this effort. The tests that were used to confirm normality of the RMSE values were the Lilliefors and Anderson-Darling tests. The results of these tests were documented in Table 2.

As noted earlier, the outputs are assumed to be independent because each run was started from scratch. Each simulation was run independently, so path sets, shortest paths, and route choices were generated without previous information. The independence of each simulation is also justified by the use of a random seed for the random number generator for each simulation process. The error of each link flow relative to a "true" value is independent of previous simulations.

The equal variance has been addressed previously with a two sample F-test for equal variance, which also passed in a majority of cases across base and impact scenario subnetwork demand. Previously, the two sample t-test for equal means only required justification for the variance to be equal between the base and impact scenarios for each subnetwork size. This proof was presented in the form of a two sample F-test for equal variance. Further proof is provided here to demonstrate equal variance across subnetwork size and impact scenario magnitude. The Bartlett and Brown-Forsythe tests may be used to determine if multiple scenarios have equal variance. Each of these tests has a null hypothesis that the variances are equal. This means that a p-value less than 0.05 for the 95% confidence level indicates the samples do not have an equal variance. Otherwise, the hypothesis of equal variance cannot be rejected.

The sample data from the ten simulations of each Guadalupe street scenario was used to test the hypothesis of equal variance across impact size. Guadalupe Street, which had ten simulations of each of the representative number of links impacted and percent capacity reductions, is the most controlled set of data for this analysis. A subnetwork size parameter of nine was used to control all variables besides the impact magnitude for Guadalupe Street. The three scenarios on Guadalupe Street follow the Goldilocks principle, accounting for the extremes and the middle with 25 percent capacity reduction to one link, a 50 percent reduction to two consecutive links, and a 100 percent capacity reduction across three consecutive links. The results of these equal variance tests and a boxplot of data are presented in Table 3 and Figure 12, respectively.

Table 3: Summary of Equal Variance Tests across Impact Scenario Magnitudes

| No. Links | Count | Mean | Std Dev |
|---|---|---|---|
| 1 | 10 | 2.54 | 0.30 |
| 2 | 10 | 2.68 | 0.20 |
| 3 | 10 | 2.95 | 0.30 |
| Pooled | 30 | 2.72 | 0.27 |

| | |
|---|---|
| Degrees of freedom | 2 |
| Bartlett's statistic | 1.6881 |
| p-value | 0.4300 |
| Brown-Forsythe statistic | 0.5479 |
| p-value | 0.5845 |

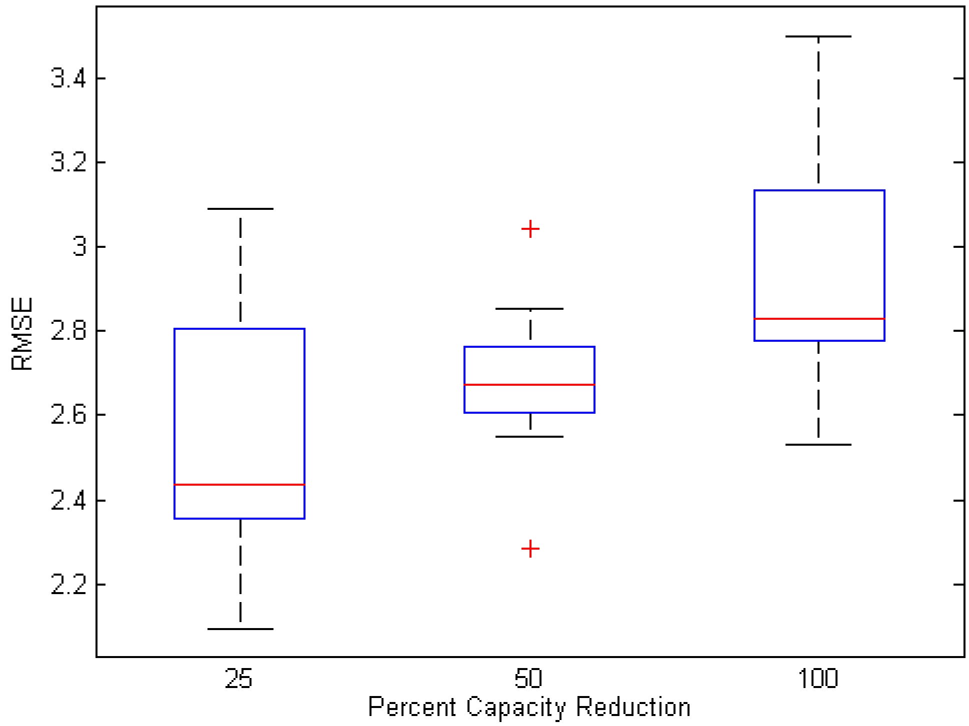



Figure 12: Boxplot of RMSE Values for the Impact Scenario Magnitude (Guadalupe Street, Subnetwork Size 9)

Table 3 indicates for both the Bartlett and Brown-Forsythe tests that the variance of the subnetwork OD matrix error is constant across impact magnitudes. The p-values were 0.43 and 0.58, which is outside the rejection region. Therefore, we assume that the null hypothesis is true and the variance is constant between impact magnitudes. The boxplot in Figure 12 graphically supports the results of the hypothesis test statistics and p-values. This boxplot also indicates the quantity of rerouting increases (captured by the RMSE) with an increase in the impact scenario magnitude. This trend follows intuition for the linear regression model, more vehicles will have to find other shortest paths as more lanes on their route are blocked.

The sample data from the common impact magnitudes was used to test this hypothesis across subnetwork size. Similar to the control method used for the equal variance test performed on impact magnitude, 50% capacity reduction scenarios were used to test for equal variance across different subnetwork sizes. These scenarios on 7$^{th}$ Street, 15$^{th}$ Street, and Guadalupe Street are the most controlled for analyzing the variance due to subnetwork size. The results of these tests and a boxplot of data are presented in Table 4 and Figure 13, respectively.

Table 4: Summary of Equal Variance Tests across Subnetwork Size Parameters

| Subnetwork Size | Count | Mean | Std Dev |
|---|---|---|---|
| 5 | 30 | 2.95 | 0.55 |
| 7 | 30 | 2.76 | 0.61 |
| 9 | 30 | 2.21 | 0.57 |
| Pooled | 90 | 2.64 | 0.58 |

| | |
|---|---|
| Degrees of freedom | 2 |
| Bartlett's statistic | 0.3083 |
| p-value | 0.8572 |
| Brown-Forsythe statistic | 0.8429 |
| p-value | 0.4339 |

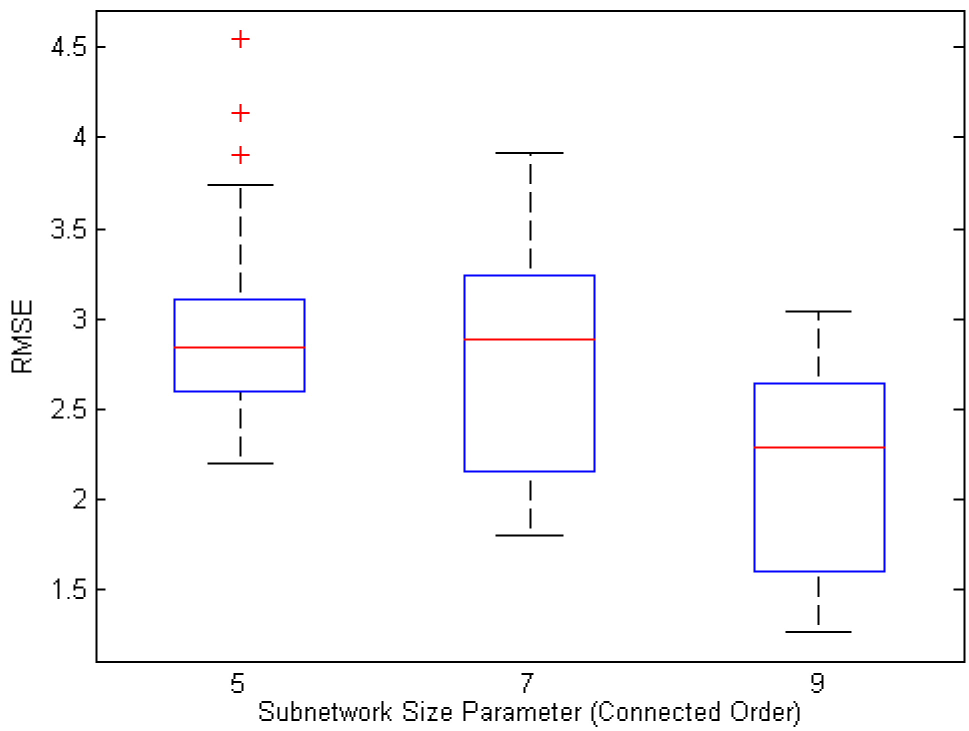



Figure 13: Boxplot of RMSE Values versus the Subnetwork Size (2 Links Impacted, 50% Capacity Reduction)

Table 4 and Figure 13 also indicate equal variance of the RMSE across subnetwork size parameters. The p-values, 0.86 and 0.43, are outside the rejection region for the null hypothesis. Again, we conclude that the null hypothesis should not be rejected and the variance of rerouting is constant with respect to subnetwork size. The boxplot in Figure 13 also indicates a decrease in rerouting as the subnetwork size increases.

These tests support the assumption of equal variance across the ranges of the independent variables, or homoscedasticity. Independence of the variables has been defended based on the randomness of the model variation. The normality of the dependent variable error has been established using normality tests. The final assumption that needs to be addressed for multiple regression is the linear relationship. Further analysis of the linear relationships between independent and dependent variables is presented in Chapter 4.3. Now that the primary assumptions for the linear regression analysis have been tested, it is important to address how to construct the model.

When formulating a predictive model, the predictor variables should be as simple as possible to maintain ease of use and interpretation. This is even more relevant for this particular use because the goal of the subnetwork creation process is to reduce the overall computational effort. The variables should at least be easy for the user to measure or lookup, and it is ideal for the values to be automatically calculated by the code creating the subnetwork. Therefore, the initial test model will include the variables that the user inputs for the impact scenario: capacity reduction, number of links, and subnetwork size parameter.

Capacity reduction can be designated through the number of closed lanes or the percentage of overall capacity eliminated. The case study was designed with percentage capacity reductions of 25, 50, and 100 regardless of the number of lanes on the given roadway. This will make it easier for a direct translation of the input variable, but it may also be a better representation of the size of the closure. If the chosen variable was the number of lanes closed, then closing a single lane on a one-lane or two-lane road would be represented with the same magnitude, but percent reduction captures the potential impact as the fraction of original capacity being reduced.

The number of links impacted can similarly be represented in two different ways. The experimental case study setup specified the number of links as represented in the GIS component of the DTA model. In this way, important connections to other features in the network, such as signalized intersections, were preserved. However, the physical length of the links may vary across different areas of the network. GIS can be used to determine the length of the actual links that are selected. This variable may have an advantage at predicting effects because it represents the actual extent of the network that is being modified. However, it may be desirable to maintain the number of links impacted as a predictor to conform to the inputs for the comparison method. Significance of these proposed variables will be evaluated with the output of different linear regression model specifications.

The subnetwork size parameter is the most critical variable for the prediction model. The size parameter, or connected order value, that has been explained previously, represents the size of the network. Incrementing the size parameter increases the extent of the subnetwork by one more link extending from the boundary. As the subnetwork size increases, more of the rerouting due to the impact should be captured. The ability of the subnetwork to capture more of the rerouting should be reflected in the decrease of the boundary error. Estimation of a coefficient that describes the linear relationship between the size parameter and the error at the boundary is what may be used to select an appropriate subnetwork size. Another value that is related to the size parameter may be the number of links included in the subnetwork. The model will be most useful if it can specifically include the size parameter, but these two options will be tested.

Within the context of this subnetwork methodology, this prediction model will provide the flexibility to estimate error for a wider range of scenarios and networks. Similar network types with different roadway characteristics or connectivity may be addressed through other variables that can be added to the model that will make it more robust. In this way, there is the potential to allow for a variable that designates a particular network type. With more data from a variety of network configurations (e.g. freeway, rural) this parameter could be calibrated. Similarly, a variable could be used for the time period of the analysis. This could provide more information for impact scenarios that are known to occur in a particular time frame, the AM peak, PM peak, or off-peak hours. The goal is to provide a robust methodology for constructing a model with respect to the data that are available.

The value of the predictive model comes from being able to determine an appropriately sized subnetwork without the need for multiple full network runs. This is made possible in the proposed model by including subnetwork size as a variable. A calibrated model will estimate the decrease in RMSE relative to an increase in subnetwork size, and different size parameters can be tested using the specified model. A strategy can then be developed for selecting the appropriate size parameter using judgment about the allowable error informed by the comparison model recommendations.

### 4.3 Specifying an Effective Model

The critical step for the linear regression approach is determining the appropriate relationship between the forecasted variable, RMSE of the induced subnetwork demand, and the predictor variables. The preliminary statistical assumptions have been addressed in Chapter 4.2, but the correlation between independent and dependent variables needs further exploration. Testing should help to indicate if the dependent variable is proportional to the predictor variables, and whether it is directly or inversely proportional. Once the proportionality is established, then it is necessary to define the nature of this proportionality: linear, logarithmic, exponential, or something else. Typical analysis for answering these questions begins with constructing and inspecting a cross tabulation of the variables and scatter plot.

The primary predictor variable required for this analysis is the size parameter, and it is expected that the RMSE decreases with an increase in subnetwork size. Cross tabulations of the boundary demand error (RMSE) ranges and the size parameters are provided in Table 5. A majority, 64.5%, of the low range of boundary demand error subnetworks were the largest subnetwork size, nine. As the RMSE range values increase there is a greater number of smaller subnetwork sizes. The largest subnetwork OD matrix error, above four, was mostly (75.0%) a subnetwork size of five and contained none of the largest subnetwork size. These cross tabulation results indicate the expected trend of decreasing RMSE for increasing subnetwork size.

Table 5: RMSE Range versus Size Parameter Cross Tabulation

| | | | Size Parameter | | |
|---|---|---|---|---|---|
| | | | 5 | 7 | 9 |
| RMSE Range | Below 2.5 | Count | 12 | 21 | 60 |
| | | % within RMSE Range | 12.9% | 22.6% | 64.5% |
| | Between 2.5 and 4 | Count | 54 | 69 | 42 |
| | | % within RMSE Range | 32.7% | 41.8% | 25.5% |
| | Above 4 | Count | 36 | 12 | 0 |
| | | % within RMSE Range | 75.0% | 25.0% | 0.0% |

Scatter plots are presented in Figure 14, Figure 15, and Figure 16 showing RMSE versus size parameter relative to different impact scenarios. The RMSE from these impact scenarios represent averaged values. These plots were used as an original test of the relationship, but have some limitations since many of the scenarios were only simulated twice. This means there is a great likelihood that the errors for these two runs may depict the extremes of the boundary demand error distribution. For instance, in Figure 14 the two runs of the 1 link scenario for 7$^{th}$ street in Figure 14 were found to have greater error than the 7$^{th}$ Street, 2-link scenario. This is not intuitive, since more impacted links should increase rerouting, but is most likely due to the 2 link, 25% capacity reduction scenarios only having two runs, while the 1 link, 25% capacity reduction scenario had ten runs.

However, the following figures provide insight to the overall trend between RMSE and subnetwork size.

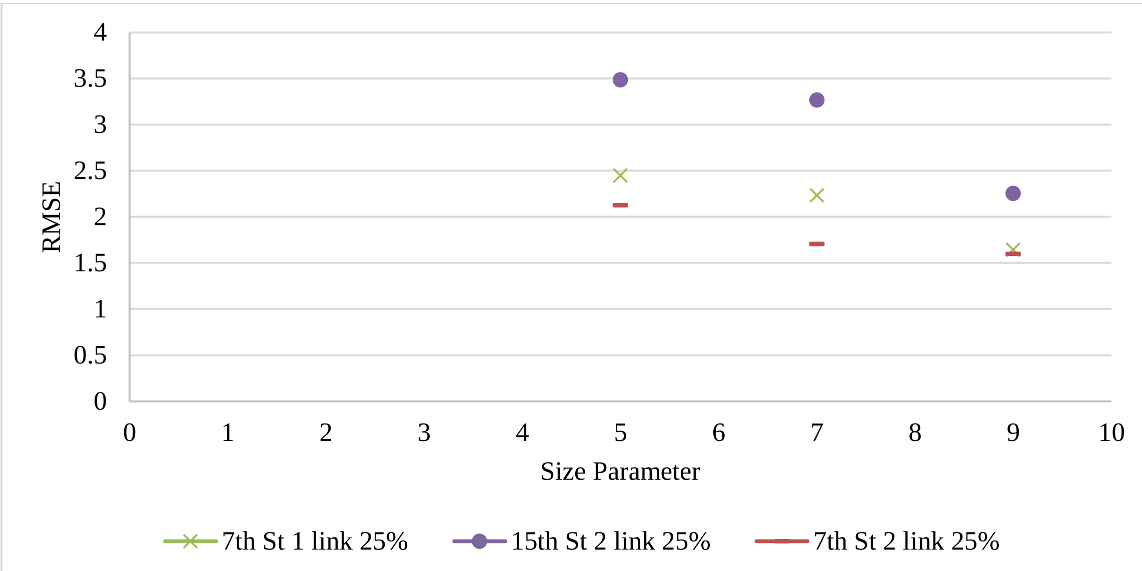



Figure 14: RMSE versus Size Parameter for 25 Percent Capacity Reduction Scenarios

Figure 15 provides more proof that the linear relationship is consistent for the 50% capacity reduction scenarios. Each of the two link scenarios in this figure were simulated ten times. In Figure 15, a similar issue can be seen for the 7$^{th}$ Street, 3 link, 50% capacity reduction scenario with a deviation from the linear trend at subnetwork size five. Again, this is potentially due to the small number of runs, two, for this scenario. The scenarios that were simulated ten times tend to appear more linear, but ten simulations is still a small sample size. While ten observations is far more likely to span the range of the distribution, there is still a possibility for capturing outliers. More than ten simulations reveals diminishing returns for capturing the average of the distribution, and this was avoided due to the computational time and effort required for a greater number of simulations. The differences between the RMSE for the 2 link, 50% capacity reduction scenarios in Figure 15 indicates that other variables may be influencing the rerouting.

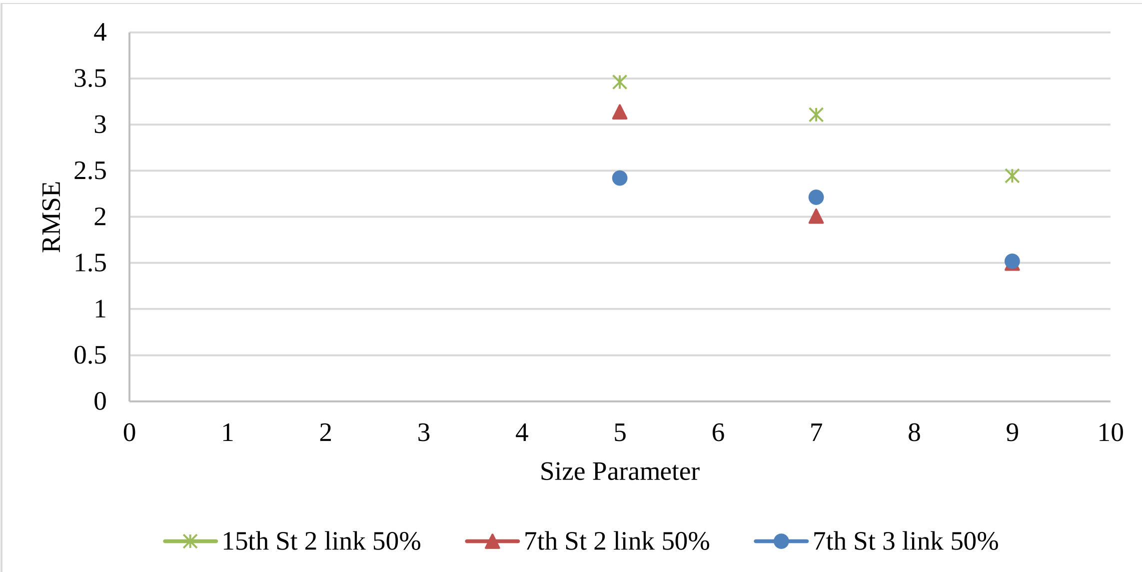



Figure 15: RMSE versus Size Parameter for 50 Percent Capacity Reduction Scenarios

Comparing Figure 15 and Figure 16, particularly the 15$^{th}$, 2 link, 25% and the 15$^{th}$, 2 link, 50% scenario indicates that there may be some interaction effect between capacity reduction and subnetwork size. This is depicted by the different trend in RMSE outcomes as subnetwork size increases with different levels of capacity reduction. This potential interaction will be investigated as more variables are added to the model.

Figure 16 reveals the linearity for the largest possible capacity reduction of 100%. In this case, the three link scenario was simulated ten times, while the one and two link were only simulated twice. Therefore, the inconsistency of the two-link scenario producing higher rerouting error at subnetwork size five and nine is probably due to the sample size. This preliminary analysis provides evidence for defending the use of linear regression, but a more thorough analysis including the multiple observations of each scenario can reveal more about the variance at the different subnetwork sizes.

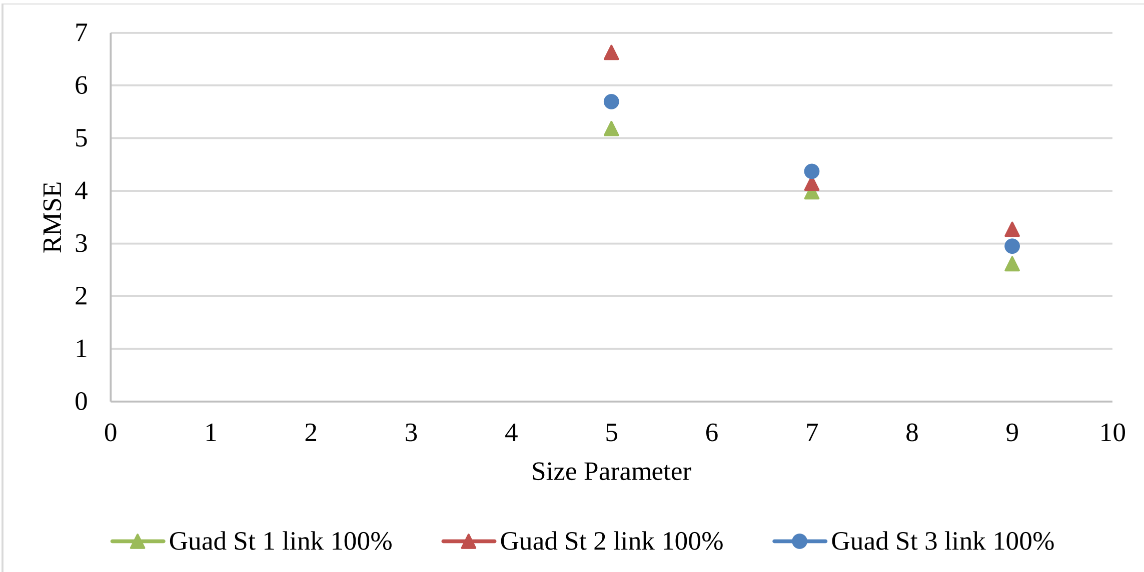



Figure 16: RMSE versus Size Parameter for 100 Percent Capacity Reduction Scenarios

Further exploration of this relationship for multiple simulations of the same scenario is presented in Figure 17, Figure 18, Figure 19. The data in the figures reaffirms that there is a linear relationship between the size parameter (connected order) and RMSE. They show that error defined at the boundaries decreases with increases in subnetwork size. As stated before, the data was found to be normally distributed and the variance (or error) appears to remain constant. These findings are corroborated with the RMSE values appearing to be evenly distributed around the trend line at each subnetwork level. Figure 17 demonstrates these linear regression assumptions for the smallest impact scenario magnitude.

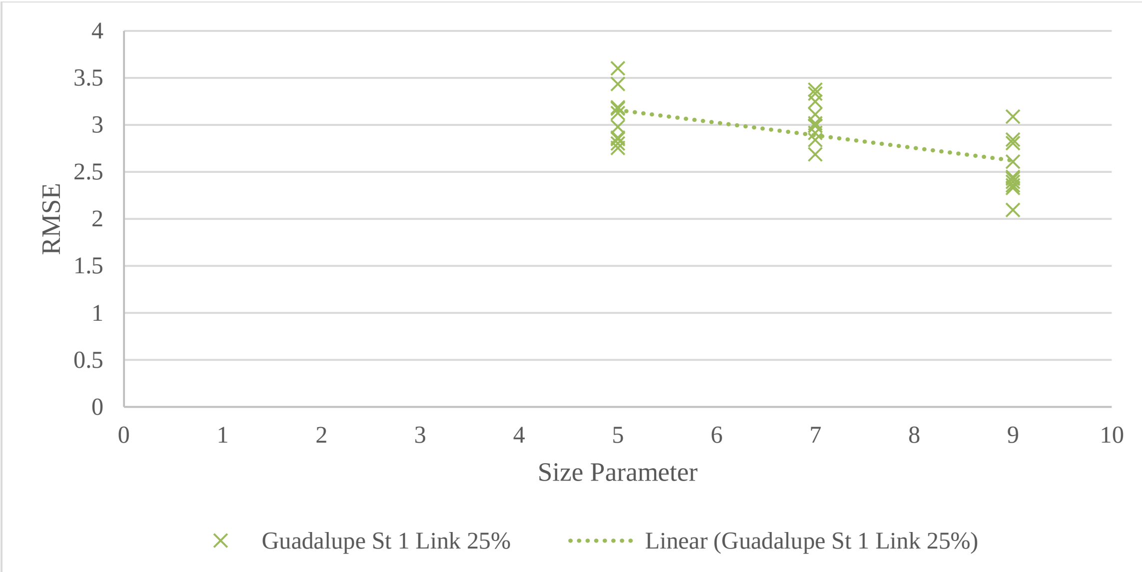



Figure 17: RMSE versus Size Parameter for Guadalupe Street, 1 Link Impacted, 25% Capacity Reduction Scenario

Figure 18 also supports the verification of the linear assumptions. The greater spread that is exhibited in the subnetwork size five is probably a result of some of the ten simulations producing outliers in the subnetwork OD matrix error. It is still clear from Figure 18 that the errors are evenly spread above and below the trend line.

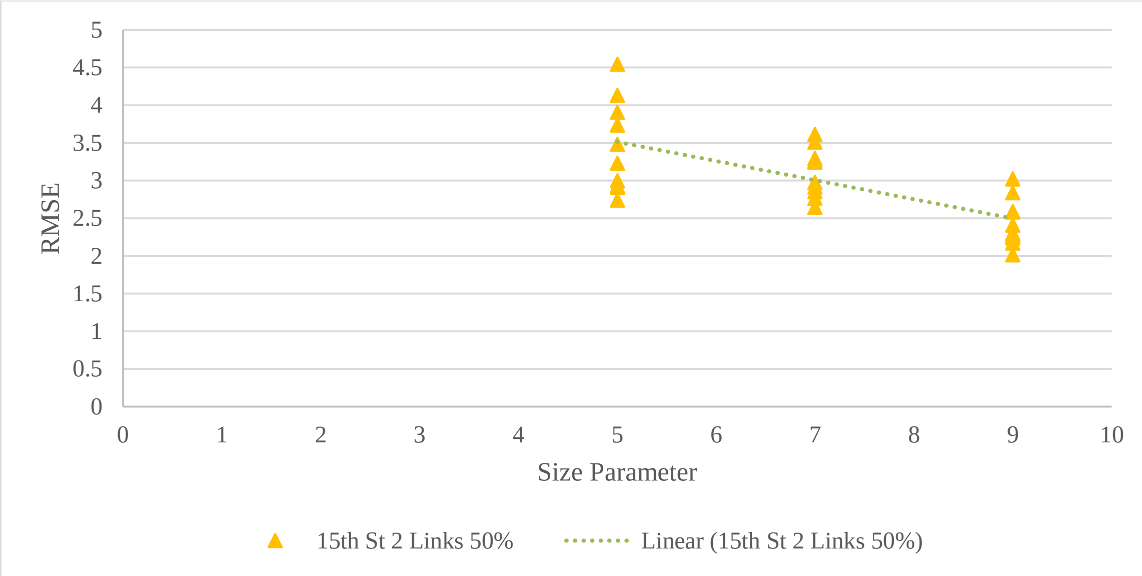



Figure 18: RMSE versus Size Parameter for 15th Street, 2 Links Impacted, 50% Capacity Reduction Scenario

Figure 19 also shows a linear relationship between RMSE and the size parameter. And again, the distribution of errors appears to fall all around the best-fit line. The differences between Figure 17, Figure 18, and Figure 19 indicate that increasing the impact scenario magnitude does increase the RMSE of the subnetwork boundary demand. Other differences, including the slope of the trend lines and the average value of the RMSE at each subnetwork size, may be the result of other variables. It is very likely that more factors than just the subnetwork size, number of links impacted, and the capacity reduction influence the amount of rerouting that occurs in a traffic network. It is desirable to develop an appropriate model for predicting the RMSE that can account for other transportation system characteristics.

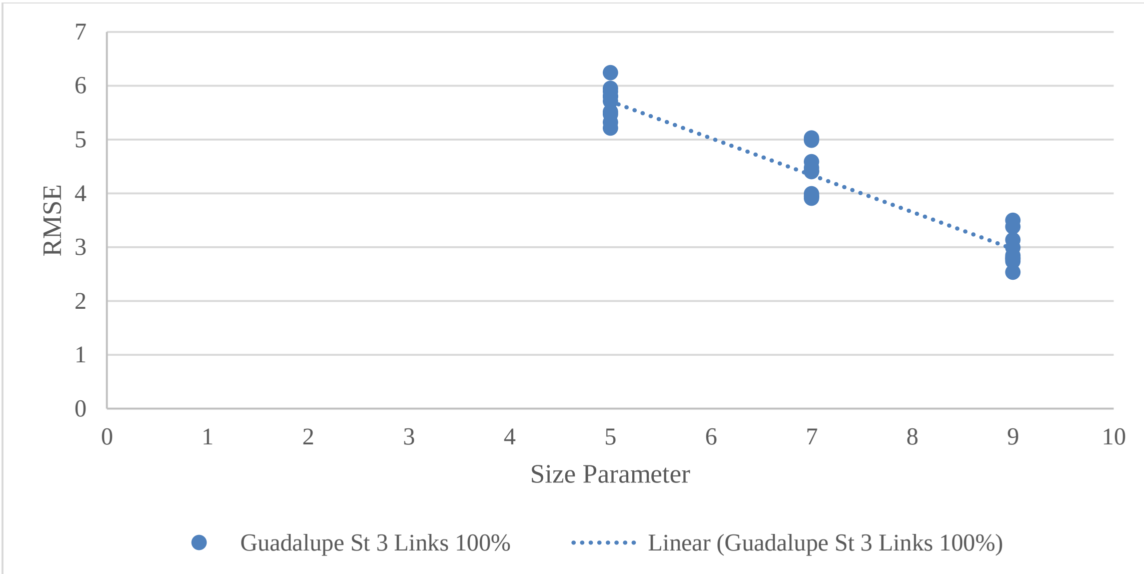



Figure 19: RMSE versus Size Parameter for Guadalupe Street, 3 Link Impacted, 100% Capacity Reduction Scenario

The spread of the RMSE values around the trend lines in Figure 17, Figure 18, and Figure 19 is further proof that there is a large amount of variation in simulation results, which needs to be addressed for this analysis (or any measure of effectiveness). The comparison test was a first step for separating the variation of the base model from the variation of the impact model. This proposed prediction model will be able to use more input variables to determine the average level of RMSE expected in the impact scenario subnetwork boundary demand.

After the linear relationship between the subnetwork size parameter and the RMSE has been established, investigating the impact scenario magnitude relationship to RMSE will address the correlation between the primary independent variables and the rerouting metric. The same type of cross tabulation and scatter plots generated for subnetwork size was created for both the number of links and the percent capacity reduction.

Table 6 reveals that the fewest number of links impacted, one, resulted in the lowest range of RMSE values. As the number of links impacted increased the proportion of scenarios within the higher error ranges increased. This trend supports the intuition that more links being impacted will induce more rerouting.

Table 6: RMSE Range versus Number of Links Impacted Cross Tabulation

| | | | Number of Links Impacted | | |
|---|---|---|---|---|---|
| | | | 1 | 2 | 3 |
| RMSE Range | Below 2.5 | Count | 9 | 2 | 1 |
| | | % within RMSE Range | 75.0% | 16.7% | 8.3% |
| | Between 2.5 and 4 | Count | 30 | 35 | 25 |
| | | % within RMSE Range | 33.3% | 38.9% | 27.8% |
| | Above 4 | Count | 3 | 5 | 16 |
| | | % within RMSE Range | 12.5% | 20.8% | 66.7% |

Table 7 reveals the same trend that is indicated in Table 6. As the reduction in capacity increases the number of scenarios that produced larger boundary demand error also increases. The cross tabulations reassure the assumptions about these independent variables, and further analysis will include the scatter plots comparing them to RMSE.

Table 7: RMSE Range versus Percent Capacity Reduction Cross Tabulation

| | | | Percent Capacity Reduction | | |
|---|---|---|---|---|---|
| | | | 25 | 50 | 100 |
| RMSE Range | Below 2.5 | Count | 7 | 4 | 1 |
| | | % within RMSE Range | 58.3% | 33.3% | 8.3% |
| | Between 2.5 and 4 | Count | 34 | 38 | 18 |
| | | % within RMSE Range | 37.8% | 42.2% | 20.0% |
| | Above 4 | Count | 1 | 0 | 23 |
| | | % within RMSE Range | 4.2% | 0.0% | 95.8% |

Figure 20, Figure 21, and Figure 22 depict the linear trend between the RMSE and both number of links impacted and the capacity reduction for the ten simulation scenarios. This is due to the Goldilocks selection method of 1 link, 25%; 2 link, 50%; and 3 link, 100% for the ten link scenarios matching the low end of capacity reduction values to the low end of links impacted, the middle with the middle, and the high with the high. This is not a problem in the overall analysis because each scenario combination was included with supplemental scenarios run twice (e.g. 2 link, 25%; 1 link, 100%; 3 link, 50%; etc.). However, for the purpose of establishing the linearity the results in Figure 20, Figure 21, and Figure 22 are sufficient.

The linear trend for the Guadalupe Street, subnetwork size five is generally represented in Figure 20. There are some inconsistencies with the 2 link, 50% scenario depicted by all of the values falling below the trend line. This is a reminder that the data from these select simulations are not perfect. However, the relationships will be tested for statistical significance and they have the potential for increasing the power of this subnetwork selection technique. The equal variance of the RMSE data points at each level of impact scenario magnitude is also visually indicated in the Figure 20 scatter plot.

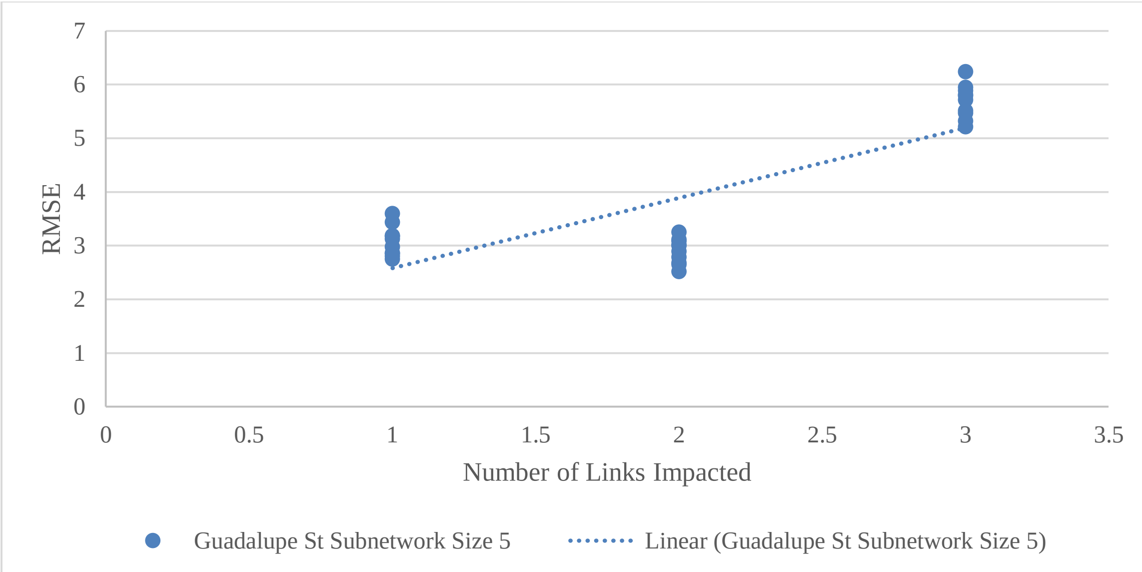



Figure 20: RMSE versus Impact Magnitude for Guadalupe Street, Subnetwork Size 5

Figure 21 also indicates a linear trend between impact magnitude and subnetwork OD matrix error. As the impact magnitude increases, the RMSE captures a greater amount of rerouting.

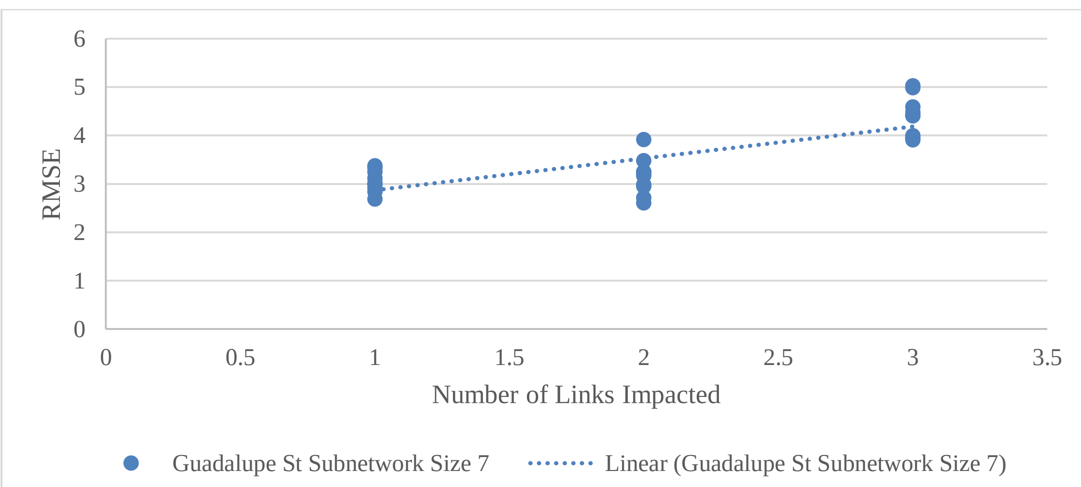



Figure 21: RMSE versus Impact Magnitude for Guadalupe Street, Subnetwork Size 7

Figure 22 depicts the same trends that are presented in Figure 20 and Figure 21. A comparison of Figure 20, Figure 21, and Figure 22 also shows the trend already established between RMSE and subnetwork size – as the subnetwork size increases the RMSE decreases. Different linear regression specifications will be used to determine appropriate predictor variables that can account for other variations observable in these scatter plots.

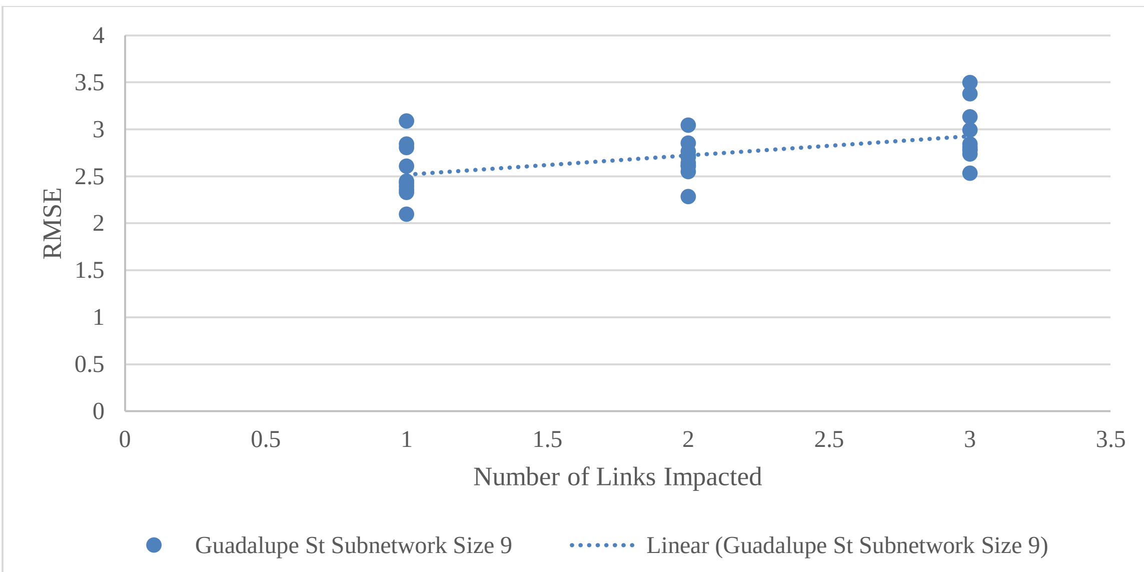



Figure 22: RMSE versus Impact Magnitude for Guadalupe Street, Subnetwork Size 9

The preliminary model construction includes the basic input variables for the impact scenario: subnetwork size, capacity reduction, and number of links impacted. This model was built using RMSE data from three locations with ten model (the Goldilocks cases) simulations for each base and impact scenario, and various other locations with two runs for each scenario accounting for all other combinations of the three primary input variables. The ordinary least squares technique was used to estimate the parameters of the linear regression model. The model equation for this initial test is represented by Equation 5. For the practical reason of using these models to predict the RMSE the error term of the linear regression equation will be left out of the equations representing the model specification.

$$RMSE = \beta_0 + \beta_1 * number\ of\ impacted\ links + \beta_2 * percent\ capacity\ reduced + \beta_3 * size\ parameter \qquad (5)$$

The data was processed in SPSS and produced the results in Table 8.

Table 8: Results of the Preliminary Linear Regression Analysis

| Model Summary | | | | | |
|---|---|---|---|---|---|
| R | R Square | Adjusted R Square | | Std. Error of the Estimate | |
| 0.685 | 0.469 | 0.464 | | 1.228 | |
| ANOVA | | | | | |
| | Sum of Squares | df | Mean Square | F | Sig. |
| Regression | 402.141 | 3 | 134.047 | 88.894 | 0.000 |
| Residual | 455.398 | 302 | 1.508 | | |
| Total | 857.538 | 305 | | | |
| Coefficients | | | | | |
| | Unstandardized Coefficients | | Standardized | | |
| | B | Std. Error | Beta | t | Sig. |
| Constant | 5.198 | 0.371 | | 14.029 | 0.000 |
| Size Parameter | -0.512 | 0.043 | -0.499 | -11.908 | 0.000 |
| Capacity Reduction | 0.026 | 0.003 | 0.456 | 9.991 | 0.000 |
| No. Links Impacted | 0.064 | 0.1 | 0.029 | 0.641 | 0.522 |

The preliminary results indicate that a little less than half the variation in the dependent variable is explained by the independent variables. The linear regression intercept is 5.2, such an RMSE value corresponds to each OD pair in each time period being off by about 5 vehicles. In other words, on average 5 vehicles were rerouted at each boundary origin in each time period. This can be quite substantial since the rerouting is probably concentrated in a particular region at the boundary. However, the intercept, 5.2, in this model is not useful as it corresponds to a subnetwork size of zero and no impact scenario, which has no meaningful interpretation. The null hypothesis for the t-statistic for each explanatory variable is that the parameter coefficient is zero. If the p-value is less than the chosen probability threshold for the hypothesis rejection region (in this case 0.05), then it indicates that the predictor is significantly correlated to the dependent variable and that its coefficient is not zero.

The size parameter and the capacity reduction were both found to be statistically significant because the t-statistics were in the rejection region. This implies that the beta coefficients for these variables are valuable predictors of their correlation with the RMSE. A preliminary interpretation of the coefficients on the capacity reduction and the size parameter relate the change in an independent variable to the dependent variable. The negative (-0.512) coefficient on the size parameter means that an increase in one of the connected order of links in the subnetwork decreases the RMSE at the boundary by half a vehicle. Another way to put this is an increase of two in the size parameter decreases the RMSE by a vehicle at each origin-destination pair and time period. The 0.026 coefficient on the capacity reduction indicates that decreasing the capacity by one percent decreases the RMSE by .02 vehicles. More intuitively, a 25 percent reduction in capacity increases the RMSE by 0.5 vehicles, a 50 percent reduction in capacity increases the RMSE by 1 vehicle, and a 100 percent reduction in capacity increases the RMSE by 2 vehicles. The number of links variable had a p-value of 0.5, which indicates it is not a significant variable in the linear model. This may indicate that other variables may be interacting with the number of links impacted and further variables are necessary to understand its effects. While the general trend of these results appears intuitive, further investigation into the specification of the model will provide more power to predictive capability.

The results of the preliminary model, informed the selection of several additional variables to be generated and tested. Alternate variables were previously mentioned in the discussion of the three primary impact scenario variables, including the length of roadway impacted was included as a surrogate for the number of links impacted and the number of links in the subnetwork was included as a substitute for the subnetwork size parameter. The number of links included in the connected order process has the potential to account for some characteristics of the surrounding network. If the network is richly connected, then the same size parameter will include more links. The other variables tested were related to the vehicular flows on the links and link characteristics. This assumes that a full model has been run and calibrated in order to establish a base flow as an input to the linear model, which is typical for DTA implementation. The base volume for each link provides the number of vehicles that are currently using that link along their path. The capacity value sets a limit on the potential increase in flow that the link can accommodate. The volume to capacity ratio is a metric for the utilization of the link in the base case. It was speculated that the volume to capacity ratio accounts for the current level of congestion in the impact scenario area and could significantly influence whether rerouting would occur and the magnitude of the rerouting. These variables also have the potential to account for directional differences in flow that may be experienced between AM and PM peak hours. Results of the different model specifications are tabulated in Table 9. The following is a description of the independent variables used.

***Key For the Model Variables***

- Size Parameter: The connected order of links from the impacted link; a variable indicating the size of the subnetwork.
- Capacity Reduction: The reduced capacity of an impacted link represented as a percent of the original capacity level.
- No. Links Impacted: The number of links with an altered capacity in the impact scenario.
- Length Impacted: The sum of the physical lengths of each impacted link.
- No. Subnet Links: The total number of links included in the subnetwork based on the size parameter used.
- Volume/1,000: The total or sum of the volumes (vehicles per hour) across the impacted link(s), scaled by dividing by 1,000 to provide a nonzero coefficient.
- Capacity/1,000: The total or sum of capacities (vehicles per hour) of each impacted link(s), scaled by dividing by 1,000 to provide a nonzero coefficient.
- Average Capacity/1,000: The sum of the capacities of each impacted link(s) divided by the number of links, scaled by dividing by 1,000 to provide a nonzero coefficient.
- Volume/Capacity: The volume to capacity ratio calculated using the computed averages.
- Size and Cap. Red. Interaction: The interaction term for subnetwork size parameter and the percent capacity reduction, calculated as the product of the two independent variables.

Table 9: Model Specifications Tested Using Link Variables

| | Estimate (t-statistic) | | | |
|---|---|---|---|---|
| Predictor Variables | A | B | C | D |
| Size Parameter | 0.971 (1.333) | **-0.512 (-11.919)** | — | **-0.512 (-11.093)** |
| Capactiy Reduction | **0.026 (9.891)** | **0.026 (11.168)** | **0.027 (11.249)** | **0.026 (-12.097)** |
| No. Links Impacted | **2.568 (2.818)** | — | — | — |
| Length Impacted | 0.000 (-0.104) | — | — | — |
| No. Subnet Links | -0.021 (-2.040) | — | **-0.007 (-10.600)** | — |
| Volume/1,000 | 0.155 (0.196) | — | — | — |
| Capacity/1,000 | **-0.371 (-2.695)** | — | — | — |
| Volume/Capacity | **-28.207 (-3.054)** | — | — | **-8.540 (-3.175)** |
| Constant | **5.318 (2.483)** | **5.292 (15.575)** | **3.772 (14.882)** | **6.707 (12.034)** |
| | Goodness of Fit and Sample Data | | | |
| Number of Cases | 306 | 306 | 306 | 306 |
| $R^2$ | 0.512 | 0.468 | 0.430 | 0.485 |
| | Estimate (t-statistic) | | | |
| Predictor Variables | E | F | G | H |
| Size Parameter | **-0.512 (-11.911)** | **-0.512 (-11.908)** | **-0.512 (-12.073)** | **-0.512 (-12.170)** |
| Capactiy Reduction | **0.026 (10.128)** | **0.026 (9.991)** | **0.026 (10.169)** | **0.026 (10.250)** |
| No. Links Impacted | — | 0.064 (0.641) | **0.548 (2.946)** | **0.719 (2.374)** |
| Length Impacted | 0.000 (0.766) | — | — | — |
| No. Subnet Links | — | — | — | — |
| Volume/1,000 | — | — | **-0.557 (-3.069)** | — |
| Capacity/1,000 | — | — | — | **-0.117 (-2.315)** |
| Volume/Capacity | — | — | — | **-9.610 (-3.537)** |
| Constant | **5.176 (13.909)** | **5.198 (14.029)** | **5.237 (14.321)** | **6.740 (11.663)** |
| | Goodness of Fit and Sample Data | | | |
| Number of Cases | 306 | 306 | 306 | 306 |
| $R^2$ | 0.469 | 0.469 | 0.485 | 0.495 |
| | Estimate (t-statistic) | | | |
| Predictor Variables | I | J | Final Impact Model | Base Model |
| Size Parameter | **-0.512 (-12.193)** | **-0.512 (-12.240)** | **0.430 (5.877)** | **-0.115 (-5.429)** |
| Capactiy Reduction | **0.026 (10.116)** | **0.026 (10.420)** | **0.135 (17.208)** | — |
| No. Links Impacted | 0.060 (0.615) | **0.717 (2.379)** | **0.719 (3.080)** | — |
| Length Impacted | — | — | — | — |
| No. Subnet Links | — | — | — | — |
| Volume/1,000 | — | **1.220 (2.103)** | — | — |
| Capacity/1,000 | — | **-0.305 (-2.973)** | **-0.117 (-3.004)** | **-0.018 (-2.415)** |
| Ave. Capacity/1,000 | -0.207 (-1.949) | — | — | — |
| Volume/Capacity | **-8.958 (-3.329)** | **-25.448 (-3.181)** | **-9.610 (-4.589)** | **-15.487 (-11.493)** |
| Size and Cap. Red. Interaction | — | — | **-0.016 (-14.353)** | — |
| Constant | **7.826 (9.243)** | **9.200 (7.061)** | 0.145 (.227) | **6.004 (20.985)** |
| | Goodness of Fit and Sample Data | | | |
| Number of Cases | 306 | 306 | 306 | 306 |
| $R^2$ | 0.492 | 0.502 | 0.701 | 0.381 |

*Bold font indicates statistical significance of that variable

## 4.4 TOWARDS PREDICTING A SUFFICIENT SUBNETWORK

Table 9 summarizes the tested model specifications that provided insight to the selection of a final RMSE prediction model. Model A represents an attempt at using all of the predictor variables and indicated that the size parameter, capacity reduction, number of links impacted, number of links in the subnetwork, scaled capacity, and volume to capacity ratio are significant variables. Models B and C indicate that there is a greater significance for the size parameter than the number of subarea links. This is a desirable result because the size parameter is a user input. Model D was an initial test of the volume to capacity ratio as a predictor and it was found to be significant, which was expected and desired from a traffic engineering perspective. Model E and F indicate that the number of links impacted performed about the same as a predictor variable as the length of roadway impacted, so the user input (number of links) will be examined further. Model F represents the preliminary linear regression model tested and showcased in Table 8. Model G reveals that using the scaled volume of the links was significant, but it was not significant when volume/capacity ratio was added to the model. This lead to removing the scaled volume variable from the model. An analysis of Table 9 indicates that Model H was a valuable specification to begin refining and investigating model outputs.

Testing multiple specifications for the regression model indicates that more variables are significant than indicated in the preliminary test. Model H, Model J, and the Final Impact Model in Table 9 incorporates the size parameter, the capacity reduction, the number of links impacted, the scaled capacity, and the volume to capacity ratio. Model I used the average capacity to address some of the collinearity between the number of links impacted and the total capacity. Although it was tested, mulitcollinearity is not a primary concern for this model for three reasons: the model is being used to predict the dependent variable, some of the collinear variables were control variables, and some of the collinear variables were products of other variables. However, Model I also reduced the significance of the number of links impacted and was rejected. Model H, Model J, and the Final Impact Model are each slight variations of one another and were all tested for their ability to predict subnetwork size. Model J includes a variable for the scaled capacity, which introduced a large amount of multicollinearity without much improvement in the coefficient of determination. The Final Impact Model was labeled as such because it was selected as the most appropriate model for predicting subnetwork size.

The final impact model is the same as Model H with the addition of the interaction term between subnetwork size and the capacity reduction identified in Chapter 4.3. Each of these variables was significant at a 95% confidence level, except for the intercept. This may be desirable since no impact scenario (zero values of all parameters) the RMSE should be zero. The initial interpretation of the coefficient for size parameter and capacity reduction are difficult to translate because the interaction term prevents a direct understanding of the individual parameter coefficients. For this model, each additional link altered in the impact scenario results in an increase in boundary error of 0.719 vehicles. A change in the base scenario capacity of 1,000 vehicles per hour of the impacted link increases the RMSE by approximately one vehicle. Also, for an increase in the original volume to capacity ratio (measured prior to the modification) of 10% there is a decrease in the subnetwork boundary error of 0.961. The interaction term has a significant impact on the prediction capabilities of the model. The coefficient of determination, or R squared, for this model was 0.701, a significant increase from 0.495 for Model H. The Final Impact Model is represented by Equation 6. The summary of the entire Final Impact Model is presented in Table 10.

.

$$RMSE = 0.430 * size\ parameter\ \ + 0.135 * percent\ capacity\ reduced + 0.719 * number\ of\ impacted\ links - 0.117 * capacity/1000 - 9.610 * volume\ to\ capacity\ ratio - 0.016 * size\ parameter * percent\ capacity\ reduced \quad (6)$$

Table 10: Results of the Final

| Model Summary | | | | | |
|---|---|---|---|---|---|
| R | R Square | Adjusted R Square | | Std. Error of the Estimate | |
| 0.837 | 0.701 | 0.695 | | 0.926 | |
| ANOVA | | | | | |
| | Sum of Squares | df | Mean Square | F | Sig. |
| Regression | 601.121 | 6 | 100.187 | 116.824 | 0.000 |
| Residual | 256.418 | 299 | 0.858 | | |
| Total | 857.538 | 305 | | | |
| Coefficients | | | | | |
| | Unstandardized Coefficients | | Standardized | | |
| | B | Std. Error | Beta | t | Sig. |
| Constant | 0.145 | 0.640 | | 0.227 | 0.821 |
| Size Parameter | 0.430 | 0.073 | 0.420 | 5.877 | 0.000 |
| Capacity Reduction | 0.135 | 0.008 | 2.408 | 17.208 | 0.000 |
| No. Links Impacted | 0.719 | 0.233 | 0.328 | 3.080 | 0.002 |
| Capacity Scaled | -0.117 | 0.039 | -0.324 | -3.004 | 0.003 |
| Volume/Capacity | -9.610 | 2.094 | -0.148 | -4.589 | 0.000 |
| Size and Cap. Red. Interaction | -0.016 | 0.001 | -2.199 | -14.353 | 0.000 |

It is intuitive that adding a link to the impact scenario will increase the rerouting experienced at the boundary. However, it might not be intuitive that an increase in the volume to capacity ratio in the base scenario would decrease the rerouting. This may be due to inability of vehicles to reroute when the volume to capacity is really high and the network is more congested. This coefficient may also be difficult to interpret due to interaction with the capacity scaled independent variable. The values in Equation 6, and the data for each scenario may be used to predict the RMSE of the subarea demand. This was done for each of the scenarios that were tested. A base model was also generated in

Table 9, with the few variables available for the base scenario. The RMSE values can be compared for a predicted base and impact scenario. This comparison was performed for Model H and Model J, but only the Final Impact Model will be included here because it performed significantly better than Model H or Model J. The results are shown in Table 11.

.

Table 11: Comparison of the Final Impact and Base Model

| Scenario | Subnetwork Size | True Base | True Impact | True Difference | Base Model | Impact Model | Model Difference | Base Model Percent Error | Impact Model Percent Error |
|---|---|---|---|---|---|---|---|---|---|
| 15th, 1, 25 | 5 | 2.84 | 2.94 | 0.10 | 3.15 | 2.36 | -0.80 | 11.0 | 20.0 |
| | 7 | 3.20 | 3.24 | 0.04 | 2.92 | 2.42 | -0.51 | 8.6 | 25.3 |
| | 9 | 2.33 | 2.28 | -0.05 | 2.69 | 2.48 | -0.22 | 15.8 | 8.8 |
| 15th, 1, 50 | 5 | 2.84 | 3.44 | 0.60 | 3.15 | 3.73 | 0.58 | 11.0 | 8.5 |
| | 7 | 3.20 | 3.85 | 0.65 | 2.92 | 2.99 | 0.07 | 8.6 | 22.4 |
| | 9 | 2.33 | 2.51 | 0.19 | 2.69 | 2.25 | -0.44 | 15.8 | 10.4 |
| 15th, 1, 100 | 5 | 2.84 | 6.89 | 4.05 | 3.15 | 6.48 | 3.33 | 11.0 | 6.0 |
| | 7 | 3.20 | 2.51 | -0.68 | 2.92 | 4.14 | 1.22 | 8.6 | 64.8 |
| | 9 | 2.33 | 2.93 | 0.60 | 2.69 | 1.80 | -0.89 | 15.8 | 38.5 |
| 15th, 2, 25 | 5 | 2.73 | 3.49 | 0.75 | 2.60 | 2.26 | -0.35 | 4.7 | 35.2 |
| | 7 | 2.92 | 3.27 | 0.35 | 2.37 | 2.32 | -0.06 | 18.6 | 29.1 |
| | 9 | 2.22 | 2.25 | 0.03 | 2.14 | 2.38 | 0.23 | 3.5 | 5.5 |
| 15th, 2, 50 | 5 | 2.73 | 3.46 | 0.73 | 2.60 | 3.63 | 1.03 | 4.7 | 5.0 |
| | 7 | 2.92 | 3.11 | 0.19 | 2.37 | 2.89 | 0.52 | 18.6 | 6.9 |
| | 9 | 2.22 | 2.44 | 0.22 | 2.14 | 2.15 | 0.01 | 3.5 | 11.9 |
| 15th, 2, 100 | 5 | 2.73 | 6.56 | 3.83 | 2.60 | 6.38 | 3.78 | 4.7 | 2.7 |
| | 7 | 2.92 | 3.85 | 0.93 | 2.37 | 4.04 | 1.67 | 18.6 | 5.0 |
| | 9 | 2.22 | 3.13 | 0.91 | 2.14 | 1.70 | -0.44 | 3.5 | 45.6 |
| 15th, 3, 25 | 5 | 2.69 | 3.42 | 0.72 | 2.43 | 2.39 | -0.04 | 9.7 | 29.9 |
| | 7 | 3.19 | 3.35 | 0.16 | 2.20 | 2.45 | 0.25 | 30.9 | 26.8 |
| | 9 | 2.23 | 2.54 | 0.31 | 1.97 | 2.51 | 0.54 | 11.4 | 0.8 |
| 15th, 3, 50 | 5 | 2.69 | 2.74 | 0.04 | 2.43 | 3.77 | 1.34 | 9.7 | 37.7 |
| | 7 | 3.19 | 3.71 | 0.52 | 2.20 | 3.03 | 0.83 | 30.9 | 18.3 |
| | 9 | 2.23 | 2.41 | 0.18 | 1.97 | 2.29 | 0.32 | 11.4 | 5.0 |
| 15th, 3, 100 | 5 | 2.69 | 6.71 | 4.02 | 2.43 | 6.52 | 4.09 | 9.7 | 2.9 |
| | 7 | 3.19 | 5.53 | 2.34 | 2.20 | 4.18 | 1.98 | 30.9 | 24.5 |
| | 9 | 2.23 | 2.76 | 0.54 | 1.97 | 1.84 | -0.13 | 11.4 | 33.4 |
| 7th, 1, 25 | 5 | 2.16 | 2.45 | 0.28 | 2.23 | 1.62 | -0.61 | 3.0 | 33.7 |
| | 7 | 1.89 | 2.23 | 0.34 | 2.00 | 1.68 | -0.32 | 5.9 | 24.6 |
| | 9 | 1.63 | 1.65 | 0.01 | 1.77 | 1.74 | -0.03 | 8.4 | 6.0 |
| 7th, 1, 50 | 5 | 2.16 | 2.36 | 0.20 | 2.23 | 3.00 | 0.77 | 3.0 | 27.0 |
| | 7 | 1.89 | 2.01 | 0.12 | 2.00 | 2.26 | 0.26 | 5.9 | 12.5 |
| | 9 | 1.63 | 1.67 | 0.04 | 1.77 | 1.52 | -0.25 | 8.4 | 9.2 |
| 7th, 1, 100 | 5 | 2.16 | 7.69 | 5.53 | 2.23 | 5.75 | 3.52 | 3.0 | 25.2 |
| | 7 | 1.89 | 2.68 | 0.79 | 2.00 | 3.41 | 1.41 | 5.9 | 27.3 |
| | 9 | 1.63 | 1.74 | 0.11 | 1.77 | 1.07 | -0.70 | 8.4 | 38.5 |
| 7th, 2, 25 | 5 | 2.09 | 2.13 | 0.04 | 2.17 | 1.67 | -0.50 | 3.8 | 21.5 |
| | 7 | 1.66 | 1.71 | 0.04 | 1.94 | 1.73 | -0.21 | 16.3 | 1.3 |
| | 9 | 1.52 | 1.60 | 0.07 | 1.71 | 1.79 | 0.08 | 11.9 | 11.9 |
| 7th, 2, 50 | 5 | 2.09 | 2.49 | 0.40 | 2.17 | 3.04 | 0.88 | 3.8 | 22.2 |
| | 7 | 1.66 | 2.01 | 0.34 | 1.94 | 2.30 | 0.37 | 16.3 | 14.8 |
| | 9 | 1.52 | 1.50 | -0.03 | 1.71 | 1.56 | -0.14 | 11.9 | 4.5 |
| 7th, 2, 100 | 5 | 2.09 | 9.29 | 7.20 | 2.17 | 5.79 | 3.63 | 3.8 | 37.6 |
| | 7 | 1.66 | 3.14 | 1.47 | 1.94 | 3.45 | 1.52 | 16.3 | 10.1 |
| | 9 | 1.52 | 1.97 | 0.44 | 1.71 | 1.11 | -0.59 | 11.9 | 43.4 |
| 7th, 3, 25 | 5 | 2.04 | 2.27 | 0.23 | 2.32 | 1.85 | -0.47 | 13.8 | 18.5 |
| | 7 | 1.88 | 2.40 | 0.52 | 2.09 | 1.91 | -0.18 | 11.3 | 20.3 |
| | 9 | 1.53 | 1.78 | 0.25 | 1.86 | 1.97 | 0.11 | 21.9 | 10.8 |
| 7th, 3, 50 | 5 | 2.04 | 2.42 | 0.38 | 2.32 | 3.22 | 0.90 | 13.8 | 33.2 |
| | 7 | 1.88 | 2.21 | 0.33 | 2.09 | 2.48 | 0.39 | 11.3 | 12.3 |
| | 9 | 1.53 | 1.52 | -0.01 | 1.86 | 1.74 | -0.12 | 21.9 | 14.8 |
| 7th, 3, 100 | 5 | 2.04 | 9.31 | 7.27 | 2.32 | 5.97 | 3.65 | 13.8 | 35.8 |
| | 7 | 1.88 | 3.03 | 1.15 | 2.09 | 3.63 | 1.54 | 11.3 | 19.9 |
| | 9 | 1.53 | 1.82 | 0.30 | 1.86 | 1.29 | -0.57 | 21.9 | 29.0 |

Table 11: Continued

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Guad, 1, 25 | 5 | 2.99 | 3.08 | 0.08 | 3.08 | 2.15 | -0.93 | 2.9 | 30.0 |
| | 7 | 2.91 | 3.05 | 0.14 | 2.85 | 2.21 | -0.64 | 1.8 | 27.4 |
| | 9 | 2.46 | 2.54 | 0.08 | 2.62 | 2.27 | -0.35 | 6.8 | 10.5 |
| Guad, 1, 50 | 5 | 2.99 | 3.51 | 0.51 | 3.08 | 3.53 | 0.45 | 2.9 | 0.6 |
| | 7 | 2.91 | 3.33 | 0.42 | 2.85 | 2.79 | -0.06 | 1.8 | 16.3 |
| | 9 | 2.46 | 2.24 | -0.22 | 2.62 | 2.05 | -0.57 | 6.8 | 8.4 |
| Guad, 1, 100 | 5 | 2.99 | 5.18 | 2.19 | 3.08 | 6.28 | 3.20 | 2.9 | 21.2 |
| | 7 | 2.91 | 3.98 | 1.08 | 2.85 | 3.94 | 1.09 | 1.8 | 1.1 |
| | 9 | 2.46 | 2.62 | 0.16 | 2.62 | 1.60 | -1.02 | 6.8 | 38.9 |
| Guad, 2, 25 | 5 | 2.89 | 3.21 | 0.32 | 2.95 | 2.16 | -0.80 | 2.3 | 32.8 |
| | 7 | 2.95 | 3.82 | 0.87 | 2.72 | 2.22 | -0.51 | 7.7 | 41.9 |
| | 9 | 2.65 | 2.93 | 0.27 | 2.49 | 2.28 | -0.22 | 6.0 | 22.2 |
| Guad, 2, 50 | 5 | 2.89 | 2.90 | 0.01 | 2.95 | 3.53 | 0.58 | 2.3 | 21.9 |
| | 7 | 2.95 | 3.15 | 0.20 | 2.72 | 2.79 | 0.07 | 7.7 | 11.5 |
| | 9 | 2.65 | 2.68 | 0.02 | 2.49 | 2.05 | -0.44 | 6.0 | 23.4 |
| Guad, 2, 100 | 5 | 2.89 | 6.63 | 3.74 | 2.95 | 6.28 | 3.33 | 2.3 | 5.2 |
| | 7 | 2.95 | 4.14 | 1.19 | 2.72 | 3.94 | 1.22 | 7.7 | 4.9 |
| | 9 | 2.65 | 3.27 | 0.62 | 2.49 | 1.60 | -0.89 | 6.0 | 51.0 |
| Guad, 3, 25 | 5 | 2.84 | 2.75 | -0.10 | 3.24 | 2.42 | -0.82 | 14.2 | 11.8 |
| | 7 | 3.45 | 3.86 | 0.41 | 3.01 | 2.48 | -0.53 | 12.5 | 35.7 |
| | 9 | 2.58 | 2.98 | 0.41 | 2.78 | 2.54 | -0.24 | 8.1 | 14.8 |
| Guad, 3, 50 | 5 | 2.84 | 2.75 | -0.09 | 3.24 | 3.80 | 0.55 | 14.2 | 37.9 |
| | 7 | 3.45 | 3.21 | -0.23 | 3.01 | 3.06 | 0.04 | 12.5 | 4.8 |
| | 9 | 2.58 | 2.63 | 0.05 | 2.78 | 2.32 | -0.47 | 8.1 | 11.8 |
| Guad, 3, 100 | 5 | 2.84 | 5.69 | 2.85 | 3.24 | 6.55 | 3.30 | 14.2 | 15.0 |
| | 7 | 3.45 | 4.37 | 0.92 | 3.01 | 4.21 | 1.19 | 12.5 | 3.7 |
| | 9 | 2.58 | 2.95 | 0.37 | 2.78 | 1.87 | -0.92 | 8.1 | 36.7 |

In Table 11 the subnetwork size that was recommended by the comparison test was highlighted. On the left side, the true impact scenario and base scenario errors are tabulated. Only the scenarios with ten simulations were highlighted with the corresponding subnetwork recommendation from the equal means tests. On the right side the base and impact model predictions are tabulated with the subnetwork recommendations from Figure 10. The errors between the model prediction and true values were also calculated. These errors averaged between 10% and 20%, but more importantly they were capable of recommending similar results to the comparison test. The results indicate that it may be possible to designate a threshold for boundary error of the impact scenario model compared to the base scenario model that can identify a sufficient subnetwork. As the predicted error from the linear regression model decreases below the predicted error of the base model, the corresponding subnetwork size is comparable to the recommended size parameter from the comparison model. These results mean that the Final Impact Model linear regression specification in Table 9 has the potential to be used as a closed form method for recommending a subnetwork size for a given network impact scenario. Correlating the results from the prediction method and the comparison method can provide reassurance of the validity of this model.

A more intuitive description of the relationship between the base scenario model and impact scenario regression model is represented in graphical form in Figure 23 through Figure 28. These graphs depict the representative scenarios that were each run with ten simulations. The lines represent the impact and base model predictions, and the arrow represents the recommended subnetwork size from the comparison method. Comparing the location of the arrow and the intersection of the impact model and base model regression lines shows the ability of the linear regression model to predict the comparison model results.

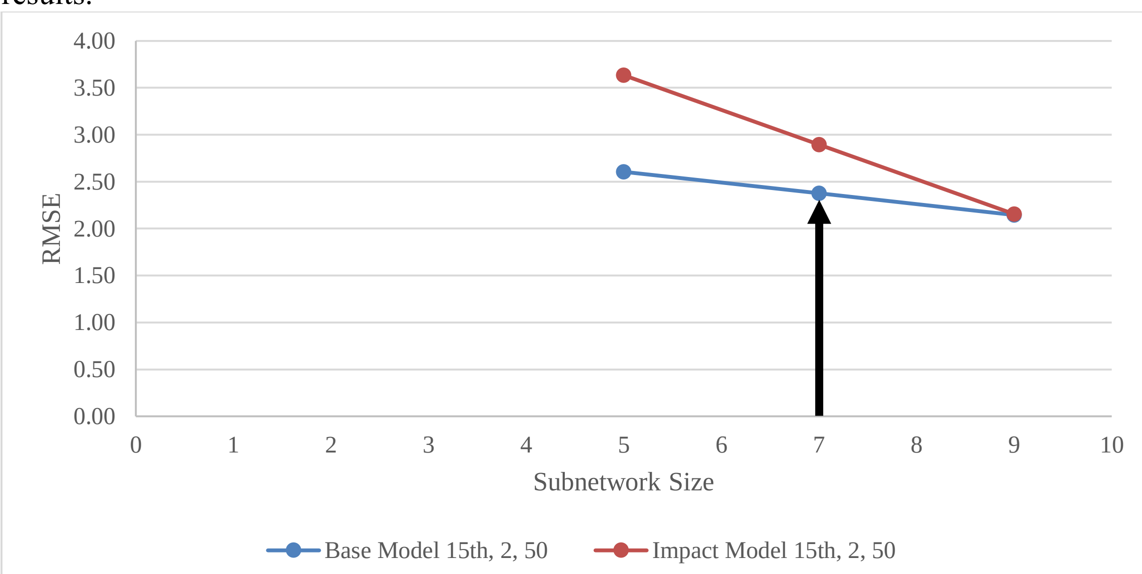



Figure 23: Subnetwork Recommendation from Prediction (Intersection) Model vs. Comparison Method (Vertical Arrow) for 15$^{th}$ Street, 2 Links, 50% Capacity Reduction Scenario

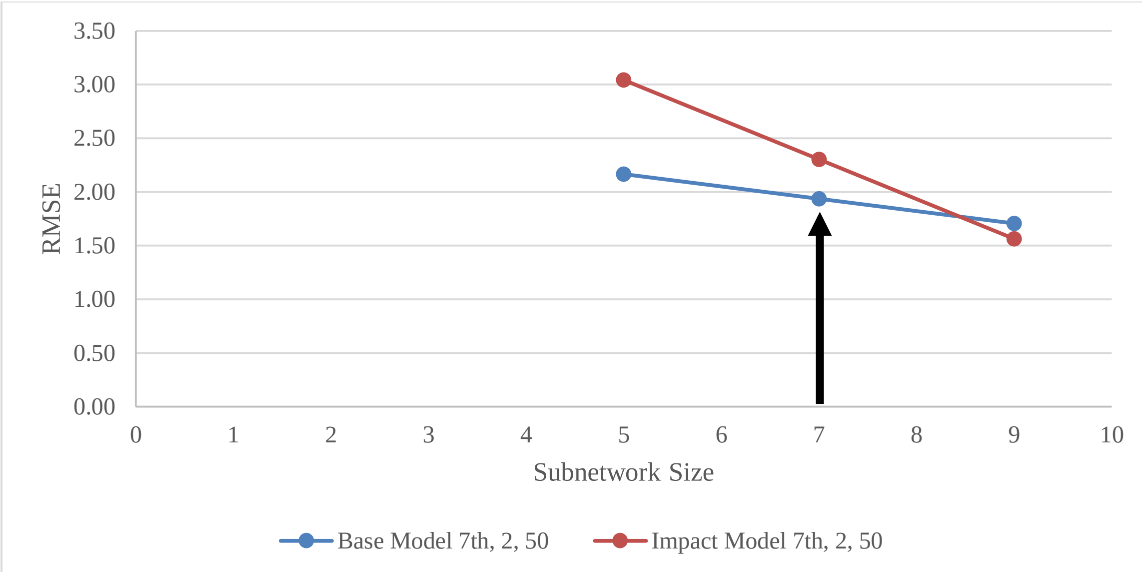



Figure 24: Subnetwork Recommendation from Prediction (Intersection) Model vs. Comparison Method (Vertical Arrow) for 7th Street, 2 Links, 50% Capacity Reduction Scenario

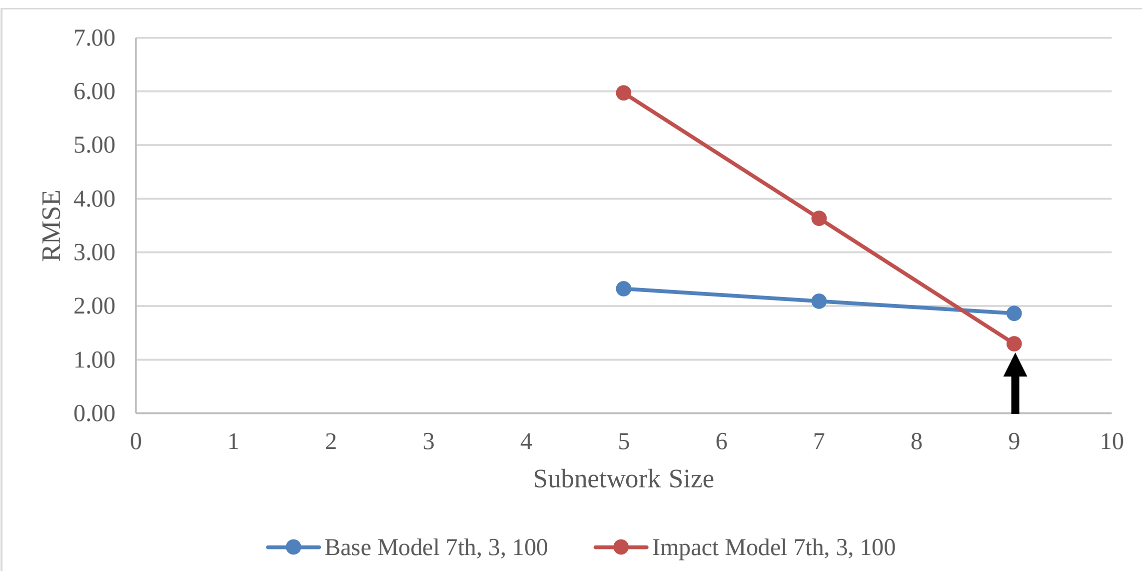



Figure 25: Subnetwork Recommendation from Prediction (Intersection) Model vs. Comparison Method (Vertical Arrow) for 7th Street, 3 Links, 100% Capacity Reduction Scenario

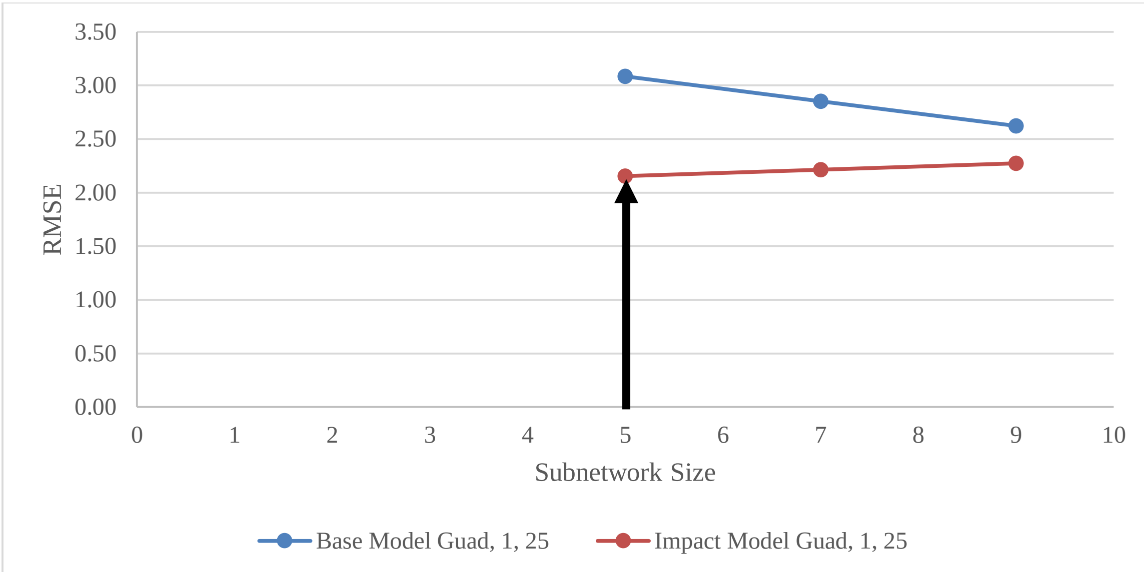



Figure 26: Subnetwork Recommendation from Prediction (Intersection) Model vs. Comparison Method (Vertical Arrow) for Guadalupe Street, 1 Links, 25% Capacity Reduction Scenario

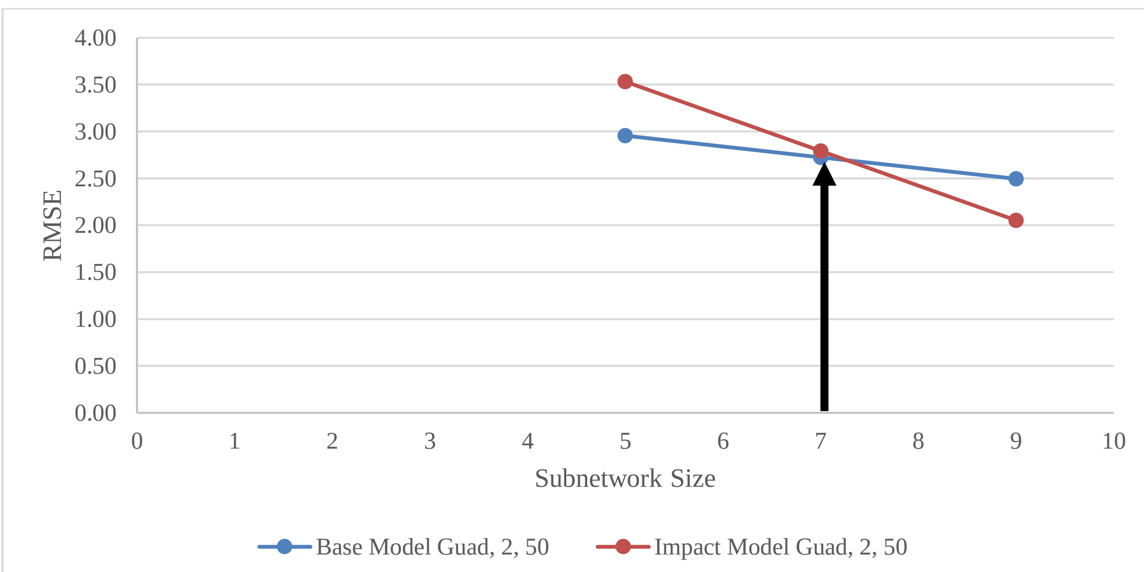



Figure 27: Subnetwork Recommendation from Prediction (Intersection) Model vs. Comparison Method (Vertical Arrow) for Guadalupe Street, 2 Links, 50% Capacity Reduction Scenario

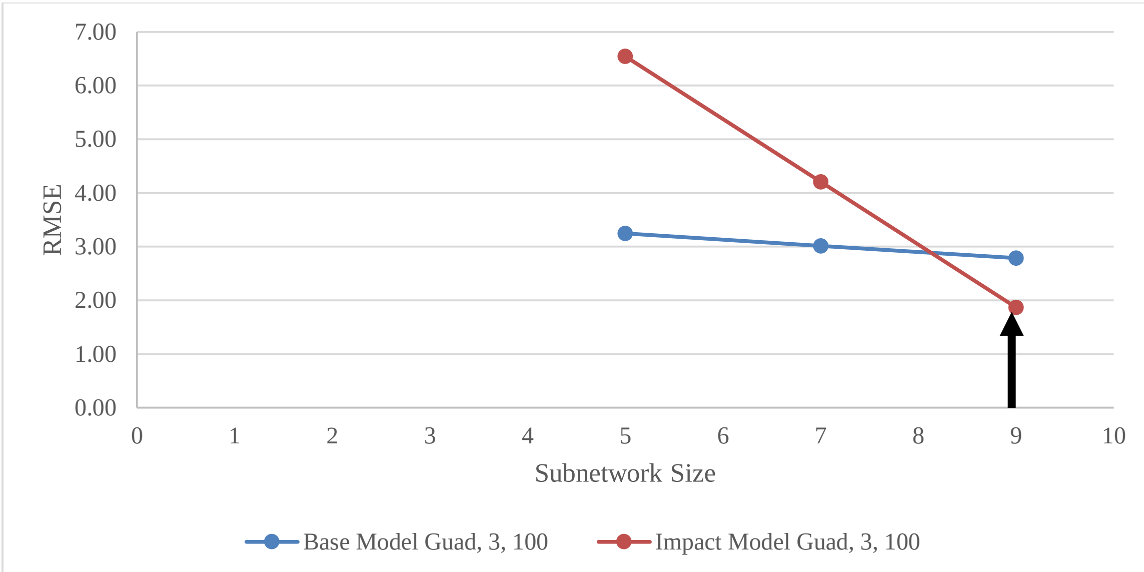



Figure 28: Subnetwork Recommendation from Prediction (Intersection) Model vs. Comparison Method (Vertical Arrow) for Guadalupe Street, 3 Links, 100% Capacity Reduction Scenario

### 4.5 Summary

This study addressed previous recommendations for a sufficient DTA subnetwork selection. Moving from average results of a statistical comparison test to a prediction model will make this procedure more robust for different scenarios and network types. The model is most useful for its ability to develop a strategy for a cutoff for subnetwork size relative to the error predicted. The original comparison process, documented in Figure 10, was used to establish a recommended subnetwork size based on tests used to determine when the base and impact models were statistically different. The results of the comparison method were used to help establish a method for using the prediction model to determine subnetwork sufficiency. Estimating the boundary error for multiple subnetwork sizes due to a specified impact can also predict the expected error for a recommended subnetwork size.

As noted earlier, given an operational RMSE predictive model, a recommended subnetwork size can be established based on what is judged to be an acceptable RMSE value relative to the base scenario. When examining the results of the comparison method with the data from the predictive model, it appears that such an RMSE threshold can be established by comparing a predicted RMSE in the base and impact scenarios. The recommendation for this linear regression prediction model is to use the first subnetwork size where the predicted impact RMSE decreases below the predicted base RMSE. Using this as the appropriate method will require only one run of the base scenario, to generate the volume and capacity parameters, to determine a sufficient subnetwork for multiple scenarios.

The contribution of this research is development of a robust method for selecting a sufficient subnetwork without the need for multiple runs of the base and impact scenarios. The major issues with the prior comparison method are the limited input variables and the time it takes to process the data. However, the comparison technique gave insight into the relationships and distributions of the data. Scatter plots indicate a linear relationship between the impact scenario characteristics and the RMSE. Statistical tests indicated that the linear regression assumptions were verified for the sample network measures. A predictive linear regression model was then used to relate subnetwork size and impact scenario characteristics to the subarea demand error. This linear equation and a similar base scenario linear equation can be used to predict the RMSE for various impact scenarios across a variety of subnetwork sizes. Finding the intersection of the predicted RMSE lines allows the user to save time in producing a more efficient subnetwork model. This methodology is presented as a flow chart in Figure 29.

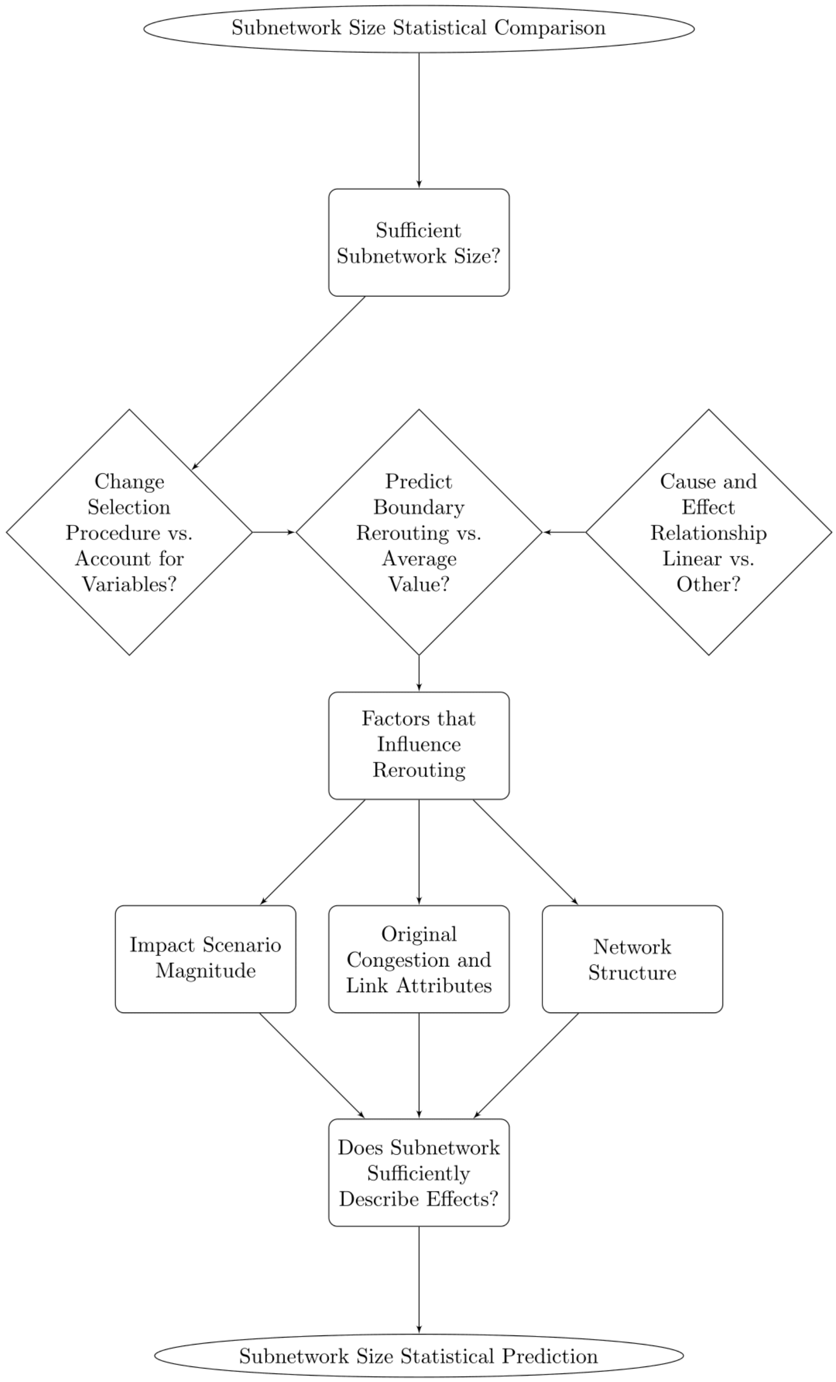



Figure 29: Flow Chart for the Subnetwork Prediction Methodology

## SECTION 2 SUMMARY: ROUTE

The route section of this dissertation describes the evolution of the tools and techniques used to solve the subnetwork selection problem. The data and experimental design for this study was used to test hypotheses for representative scenarios. This subnetwork analysis can be divided into two major steps: a statistical comparison test and a statistical prediction model. These developments are a substantial contribution because they establish methods for evaluating the effects of spatial extent on subarea accuracy, identify characteristics that influence spatiotemporal congestion in dynamic traffic assignment, and expose the limitations of certain metrics for subnetwork applications.

The methodology for subnetwork analysis was whittled down from the common methodologies within transportation systems analysis. This section identified the appropriate analysis tools (analogous to path generation), tested their effectiveness (calculated shortest paths), and determined the most efficient method to pursue (path assignment). Route generation, shortest route identification, and altering route assignments are the three components of routing in DTA and are analogous to the strategy used to develop this methodology, or the strategy for developing any new procedure. This study attempted several methods in the following order: 1) link-based statistics, 2) static traffic assignment subarea boundary demand adjustment, 3) logit (utility) formulation for boundary demand adjustment, 4) distance measure of furthest extent of rerouting, 5) percent displacement of vehicles at the boundary for users on paths containing the primary impacted link(s), 6) standard subnetwork selection technique (connected order), 7) induced, dynamic subarea origin-destination demand comparison, 8) altered connected order subnetwork selection techniques for different types of network structure, 9) prediction of the impact scenario subnetwork OD error, 10) selection of a subnetwork at the intersection of regression lines for base and impact scenario subnetwork demand. The final linear regression analysis was chosen as the optimal method because it relieves the user from the burden of running the network multiple times, accounts for other network details like volume to capacity (important for addressing base scenario network congestion), and can be easily automated. The intersection point of the regression lines finds when the impact and base subnetwork boundary demand are statistically similar, just like the comparison model. At this subnetwork size the boundary rerouting in the impact scenario is no greater than the base scenario and the subarea inputs are sufficient; therefore, the subnetwork size is sufficient.

The transferability of the proposed methods is the greatest issue that must be addressed prior to implementation. Section 3 will elaborate on the results of the case study, including an additional test network to help define the recommendations associated with this procedure. A summary of the subnetwork models and their performance will help to address the scope of these techniques. Also, an attempt to develop a quantitative metric for classifying the dynamic traffic assignment network structure will be introduced. Finally, a conclusion will briefly summarize the contribution of this dissertation.

# SECTION 3: DESTINATION

The destination section of this dissertation describes how the methods proposed in the route section may be validated through travel time analysis, transferred to other types of networks, and implemented in practice. It also aims to further define the scope of the contribution and where further research is needed. The findings of this dissertation are not limited to the specific application of subnetwork studies, but reveal many issues with the general dynamic traffic assignment process. The methodology proposed here adds one more step forward in the progress on multi-resolution analysis.

In transportation planning, the destination is analyzed for the demand needs associated with the facilities located in the traffic analysis zone. For this dissertation, the destination section will address the needs of those who will actually perform the traffic analysis. The destination section focuses on the desired goals of this dissertation and how they were accomplished. It also represents how the methodology in Chapter 4 and Chapter 5 develops the state of subnetwork research. Section 1 introduced the current state of subnetworks (origin), Section 2 establishes a new means for addressing the subarea selection need (route), and Section 3 will describe where the state of the art now stands (destination).

# Chapter 5: Case Study Analysis

The methodology in this dissertation presents a novel metric for identifying the spatial scale to which traffic flow changes may propagate due to a capacity perturbation in the network. This metric focuses on the demand input to the subnetwork region, and the incremental time required to simulate larger subnetworks. Analyzing the induced subarea demand allows for an aggregate measure of the rerouting that occurs at the boundary. While this is the most appropriate metric for determining sufficient subnetwork size, the best measure of effectiveness of a candidate subnetwork is travel time.

Travel time can be directly related to user cost and it is the most desirable traffic simulation output for multiple reasons. The primary reason is the underlying assumption of user equilibrium – every vehicle is attempting to complete its trip in minimal time. In DTA, travel time is also a more reliable link performance metric than flow because at low flow conditions, small increases in link flow may cause no change in travel time (free flow speeds remain and users are not affected). Although, density or possibly speeds are typically used for characterizing levels of service, in construction work zone scenarios where the traffic engineers' goal is to minimize user cost, the most effective way of measuring that is by measuring travel times. Therefore, accurately predicting the travel time, in other words the delay or impedance to the user, is a primary concern of this analysis. Travel time may also be implicitly derived from a fundamental flow parameter, speed, for a given link length.

This chapter will analyze the case study data from the Austin subnetwork procedure and an alternate case for the Dallas regional network. Results from these case studies will then be used to generate qualitative and quantitative recommendations for users that wish to implement this subnetwork technique.

## 5.1 EVALUATION OF THE AUSTIN SUBNETWORK CASE STUDY RESULTS

The model validation for this analysis will assume that if the recommended subnetwork size produces reliable travel times, then the model is effective. Travel times will never remain constant from simulation to simulation, regardless of the size of the network, due to the innate variation in the model algorithms. However, it is anticipated that a high level of error in travel times is likely if a subnetwork is too small and the error should decrease with increasing subnetwork size. It is desirable for the linear regression model process to recommend a size that minimizes the error in the travel time output.

This validation step aims to determine the ability of a subnetwork to predict measures of effectiveness with respect to the full network. Errors in the link travel times predicted by the different models will be compared across the range of tested subnetwork sizes. In order to derive the error in the subnetwork travel time predictions without geographic bias it is important to compare the same set of links. In other words, the data should be aggregated over the same spatial extent. For this analysis the full network simulation of the impact scenario will be treated as the true value for travel times. Then, each link's travel time in the full network will be compared to the travel times generated by subnetworks of connected order size five, seven, and nine. Since the smaller subnetworks are a subset of the larger subnetworks, the links in the smallest subnetwork size, five, will be used for comparison. The root mean square error of the travel times will be calculated on a link by link basis (for the most reliable simulation data – the peak hour), and then averaged to produce an overall travel time error for all links in the connected order of five.

Table 12 documents the results of the travel time analysis for the Austin subnetwork case study. Generally, the travel time error decreases with an increase in subnetwork size. It also appears that for each scenario there is a point of significant decrease in the error.

For almost all network sizes this decrease occurs at or below the recommended subnetwork size from the regression model. This is a desirable feature of the subnetwork selection method based on the linear regression model. By slightly overestimating the size of subnetwork needed it guarantees that the accuracy of the subnetwork model will be maintained. The time savings of a smaller network would improve efficiency, but is not as important as the accuracy of the predictions.

Table 12: Travel Time Validation of Austin Subnetwork Recommendations, RMSE of Travel Times on Each Shared Link (Seconds)

| Scenario | Five | Seven | Nine |
|---|---|---|---|
| 7th, 1, 50 | 9.37 | 3.98 | 3.43 |
| 7th, 2, 50 | 6.13 | 4.06 | 3.41 |
| 7th, 3, 50 | 5.87 | 3.93 | 3.52 |
| 7th, 1, 100 | 8.53 | 4.31 | 4.33 |
| 7th, 2, 100 | 8.60 | 3.05 | 4.14 |
| 7th, 3, 100 | 8.60 | 3.34 | 3.49 |
| 15th, 1, 50 | 10.19 | 7.87 | 10.71 |
| 15th, 2, 50 | 8.26 | 46.65 | 6.97 |
| 15th, 3, 50 | 8.84 | 43.19 | 7.16 |
| 15th, 1, 100 | 11.86 | 6.97 | 7.46 |
| 15th, 2, 100 | 7.26 | 8.27 | 8.73 |
| 15th, 3, 100 | 17.20 | 32.17 | 14.00 |
| Guad, 1, 25 | 4.70 | 3.07 | 3.36 |
| Guad, 1, 50 | 4.48 | 2.95 | 3.07 |
| Guad, 2, 50 | 8.55 | 8.41 | 8.60 |
| Guad, 3, 50 | 4.51 | 3.01 | 3.04 |
| Guad, 1, 100 | 4.81 | 3.07 | 3.24 |
| Guad, 2, 100 | 4.74 | 3.00 | 2.95 |
| Guad, 3, 100 | 9.45 | 9.73 | 9.67 |

The magnitude of the error for the Austin subnetwork is generally between five seconds and one minute. This may not seem like a substantial variation in travel time, but if this variation is added over dozens of links within a path, it could potentially result in calculating path travel times that significantly deviate from the true, expected value. Therefore, it is essential to apply this methodology and reduce (or understand) the error in subarea performance metrics. The RMSE of the induced subarea boundary demand error provides a measure of the reliability of input data for the subnetwork. The subnetwork size recommendations of the predictive models are likely to overestimate the network size needed to generate accurate results, as determined from Table 12. This assertion follows the assumption that for a given induced demand matrix the subarea should be able to replicate the output data (travel time) of the full network under the same scenario conditions. Therefore, minimal differences in the subnetwork OD matrix, detected by the regression models, may indicate a need for a larger subnetwork, while a smaller subnetwork may have been capable of producing reliable travel times. This travel time analysis is the desired proof of the efficiency of the linear regression model for informing subarea selection.

In order to provide some intuition about the characteristics of the subnetwork, several graphs are presented to explain the influence of spatial extent on: simulation time, relative gap, the number of vehicles, the number of links, and the number of nodes. The simulation time is represented by the DTA algorithm computation time only, so it does not include user processing. The relative gap is a convergence criteria for network traffic assignment, which is calculated using Equation 7:

$$Relative\ Gap = \frac{Total\ System\ Travel\ Time\ of\ Vehicles\ on\ Current\ Path}{Total\ System\ Travel\ Time\ of\ Vehicles\ on\ Shortest\ Path} - 1 \qquad (7)$$

The number of vehicles is the total count of all vehicles that traverse the subnetwork. Finally, the number of links and nodes is the count of links and nodes included in the connected order of the subnetwork. The number of centroids and connectors follow a similar trend as the number of links and is not presented here.

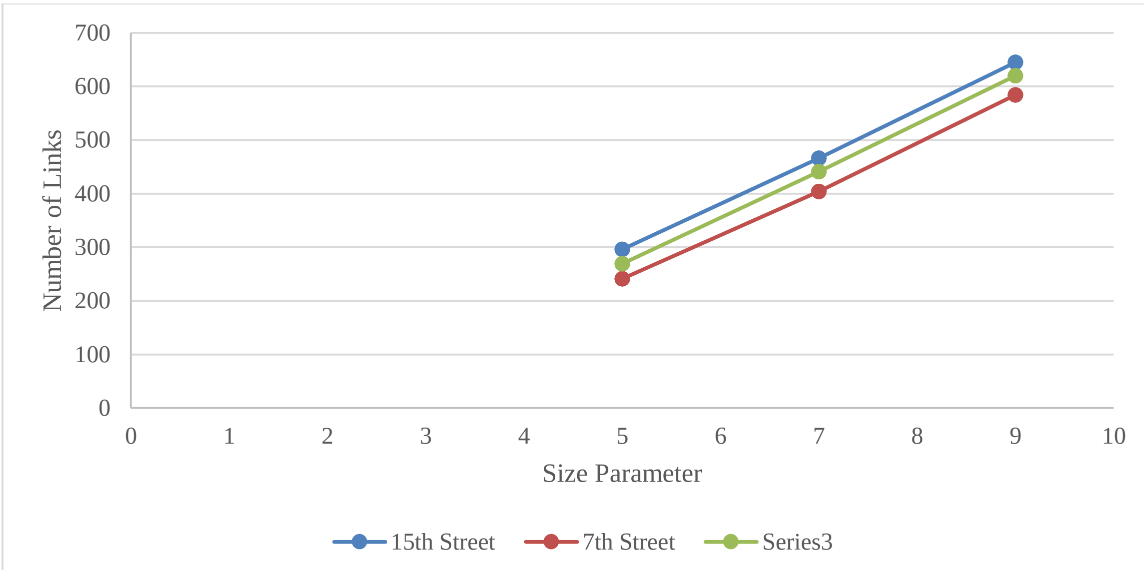



Figure 30: Summary of the Number of Links versus Subnetwork Size, Austin

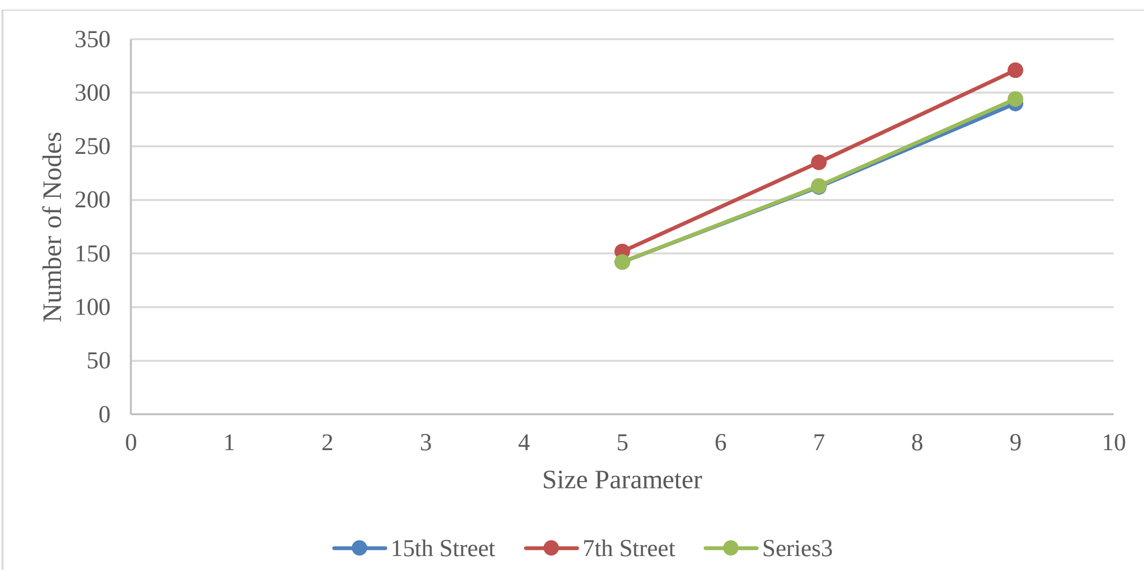



Figure 31: Summary of the Number of Links versus Subnetwork Size, Austin

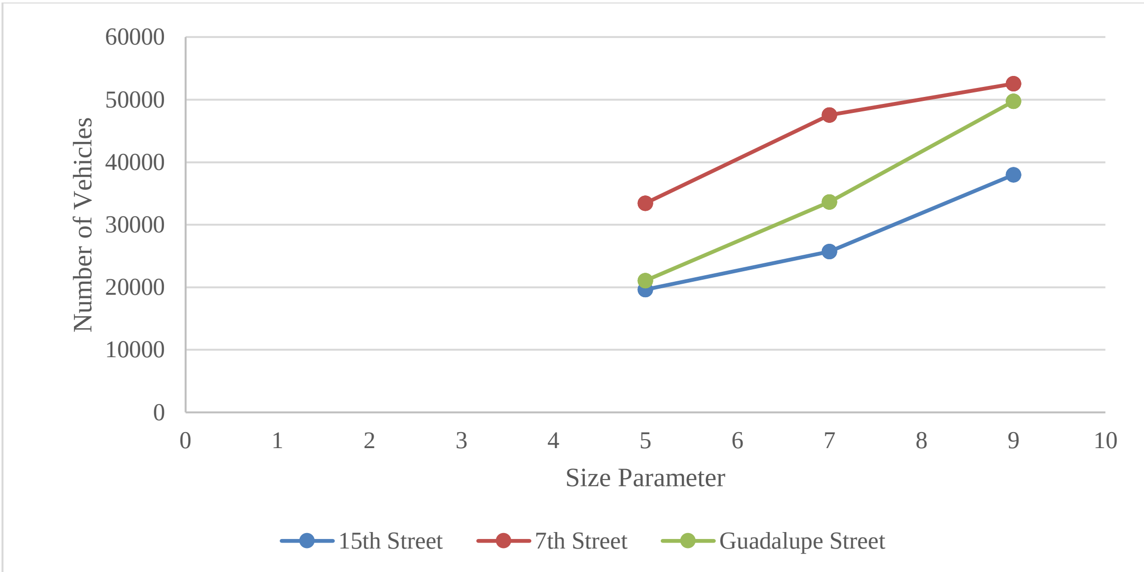



Figure 32: Summary of the Number of Vehicles versus Subnetwork Size, Austin


Figure 33: Summary of the Simulation Time (Hours) versus Subnetwork Size, Austin


Figure 34: Summary of the Relative Gap versus Subnetwork Size, Austin

Table 13: Representative Characteristics for Full Network versus Subnetwork Characteristics, Austin

| Spatial Extent | Number of Vehicles | Number of Links | Number of Nodes | Simulation Time (Hours) | Relative Gap |
|---|---|---|---|---|---|
| Full | 97606 | 1578 | 717 | 2.37 | 1.39 |
| Nine | 37988 | 645 | 290 | 0.83 | 0.86 |
| Seven | 25710 | 466 | 212 | 0.65 | 0.22 |
| Five | 19628 | 296 | 142 | 0.55 | 0.17 |

Table 14: Representative Percent of Full Network Characteristics for Subnetworks, Austin

| Spatial Extent | Number of Vehicles | Number of Links | Number of Nodes | Simulation Time (%) | Relative Gap |
|---|---|---|---|---|---|
| Full | 100% | 100% | 100% | 100% | 100% |
| Nine | 39% | 41% | 40% | 35% | 62% |
| Seven | 26% | 30% | 30% | 28% | 16% |
| Five | 20% | 19% | 20% | 23% | 12% |

It can be seen in Figure 30 through Figure 34, and in Table 13 and Table 14 subnetwork analysis offers significant reductions in computation time and a valuable improvement in convergence of the DTA model. The number of nodes and links have a linearly increasing relationship with subnetwork size. Improvements in the relative gap and simulation time as subnetwork size deceases is expected and desirable. Although, a majority of the percentage decrease occurs at the first subnetwork size of nine, the incremental improvements at sizes below nine are critical to reducing simulation time. It is likely that multiple scenarios with two to ten simulations will be necessary for a thorough traffic analysis. The savings of a half to one hour for simulation time could mean a saved half work day for practitioners.

### 5.2 Dallas Subnetwork Cross Case Study

A cross case study was included to test the transferability of the proposed method. This methodology and the trends identified in this dissertation are intended to provide insight into all types of traffic network analysis and help other researchers understand the key factors in spatial resolution for transportation modeling. A Dallas downtown model was generated to test aspects of how this methodology may be applied in other regions and scenarios.

The downtown Dallas model, which is analogous to the downtown Austin model, was generated from the Dallas-Fort Worth regional model. The 2007 network model was created at the Center for Transportation Research. An extensive revision of the network attributes and topology attempted to update the data to match the newly available 2012 regional origin-destination demand matrix. The current model was calibrated with the best available Dallas traffic counts at a number of critical locations in the downtown area. The DTA simulation output was typically within 20-30% of the real world count volumes.

Greater deviations between traffic counts and model results likely occurred due to inflated capacities at particular locations. This may be due to an over generalization of roadway capacity in the regional model or an aggregate representation of links in certain portions of the network. The downtown Dallas model was also not to a typical level of convergence desired for a DTA network, but was acceptable considering the network size.

Table 15: Downtown Dallas DTA Model Validation with Traffic Counts

| Roadway | Reference Street | Direction | Number of Lanes | Traffic Count | DTA Volume Output | Percent Error |
|---|---|---|---|---|---|---|
| IH 345 | Pacific Ave | NB | 5 | 6483 | 6126 | 6% |
| IH 345 | Pacific Ave | SB | 5 | 4267 | 3965 | 7% |
| SP 366 | Lamar St | EB | 4 | 5169 | 4101 | 21% |
| SP 366 | Griffin St | WB | 4 | 3089 | 2463 | 20% |
| IH 35E | Woodall Rodgers Fwy | NB | 5 | 1845 | 6734 | 265% |
| IH 35E | Woodall Rodgers Fwy | SB | 5 | 5245 | 4097 | 22% |
| IH 30 | Jefferson Blvd | EB | 4 | 3365 | 4365 | 30% |
| IH 30 | Jefferson Blvd | WB | 4 | 4040 | 4902 | 21% |
| IH 30 | Akard St | EB | 3 | 4200 | 5148 | 23% |
| IH 30 | Akard St | WB | 3 | 3793 | 6472 | 71% |
| IH 30 | 2nd Ave | EB | 4 | 5522 | 4366 | 21% |
| IH 30 | 2nd Ave | WB | 4 | 4327 | 5842 | 35% |
| IH 35E | Hi Line Dr | NB | 5 | 5948 | 10901 | 83% |
| IH 35E | Hi Line Dr | SB | 5 | 6770 | 6119 | 10% |
| IH 30 | Beckley Ave | EB | 3 | 4355 | 5985 | 37% |
| IH 30 | Beckley Ave | WB | 3 | 3865 | 4209 | 9% |
| IH 35E | Colorado Blvd | NB | 5 | 5759 | 6729 | 17% |
| IH 35E | Colorado Blvd | SB | 4 | 4108 | 2617 | 36% |

The Dallas subnetworks were selected in a similar manner to the Austin subnetworks. The same ranges were used for the number of links impacted, percent capacity reduction, and subnetwork size. The representative scenario locations were north bound Pearl Street, a three lane north-south arterial; west bound Jackson Street, a three lane one-way arterial; and north bound Woodall Rodgers Freeway, a four lane freeway on the northwest side of the downtown loop roadway. Link capacity perturbations examined included 1 Link, 25%; 2 Links, 50%; and 3 links, 100% for Jackson Street and Pearl Street. A 25% capacity reduction on one link was simulated on Woodall Rodgers Freeway.

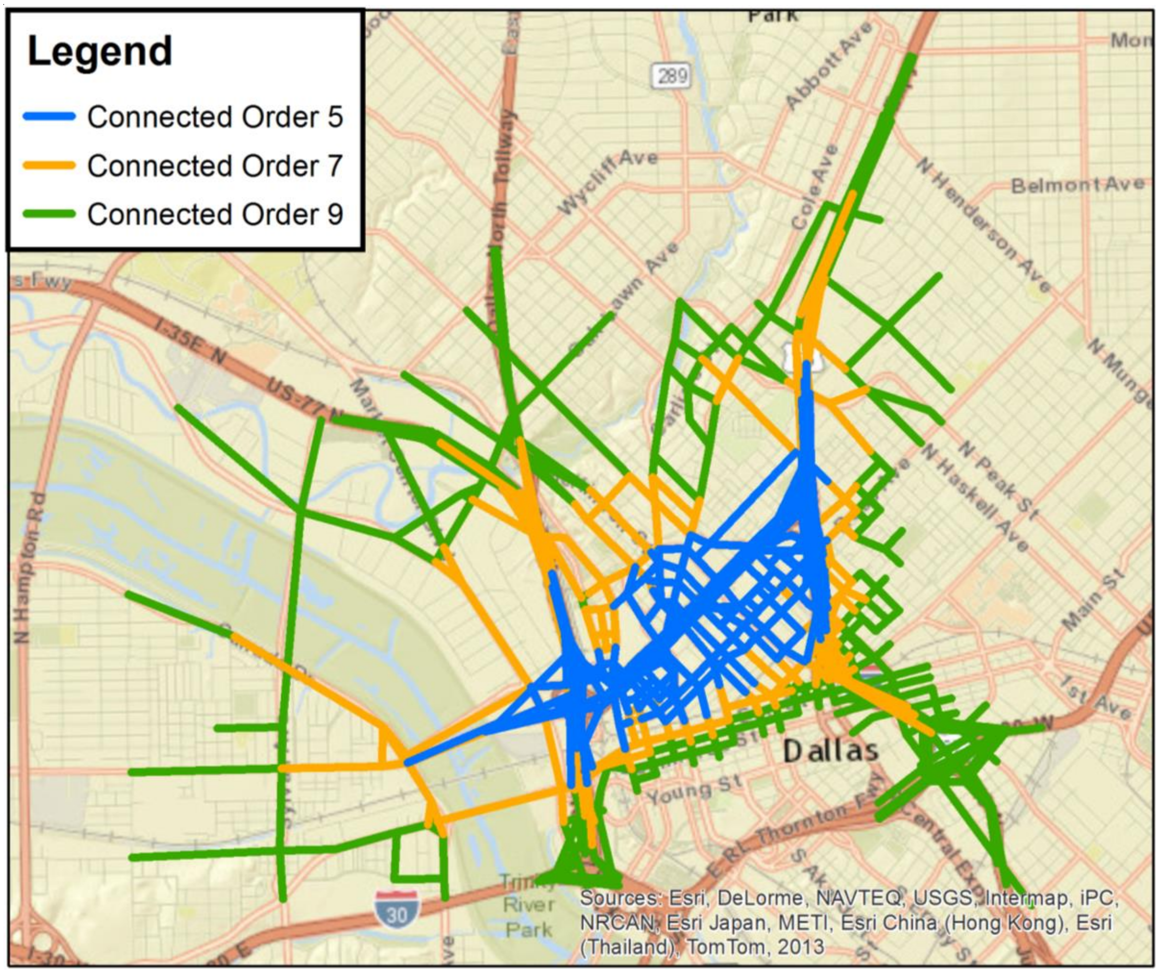



Figure 35: Example Subnetwork Connected Order Selection for Woodall Rodgers Freeway Impact Scenario, Dallas

There are several differences between the Dallas and Austin regional network and subnetwork analyses. Each of the Dallas networks were only simulated once. The Austin networks were either run ten times or twice, which allowed the capture of the within base scenario rerouting. Using a fixed random number seed to control variability, only one run of the Dallas network was tested to examine the exact difference between the base and impact scenarios. This allows for the opportunity to isolate changes in travel time and induced subarea demand as a result of the impact scenario.

The Dallas network was simulated in a slightly different manner than the Austin network. For the Dallas regional network a time step of 6 seconds was used to make the simulation of the full network manageable. This 6 second time step was maintained for consistency across the Dallas networks. Another major difference is time period of the demand data for the models – the Dallas model uses AM peak period demand and the Austin model uses PM peak period demand. The Dallas regional demand data also included trucks in addition to passenger cars, while the calibrated Austin network only had passenger cars. The routing of trucks is identical to routing for passenger cars; although, the length of the trucks is a factor in the congestion on certain links. Also, the prevalence of trucks varies throughout areas in the network – the greatest concentration of trucks occurs on the freeway. The downtown Dallas model required a total of 90 iterations to reach an acceptable level of convergence, compared to the 50 iterations used for the Austin network.

Several issues may have led to the poor convergence of the Dallas model. Primarily, the 2007 network was updated to match recent changes to the Dallas regional roadways. However, this process focused mostly on downtown freeways, so less significant improvements may not be reflected in the updated network. This means that there is likely an overestimation of congestion in the network simulation. The worse convergence, or higher relative gap, is also a result of less calibration for the Dallas model compared to the downtown Austin network. The downtown Austin model is a smaller region and had an extensive review process of all link attributes. The detail in the downtown Dallas network is not as complete as Austin and several minor streets are not included in the network. This has the potential to create problems for convergence because alternative routes are limited relative to reality.

The results of calculating the RMSE of the induced subarea boundary demand are organized in Table 16. The true impact column tabulates the one to one comparison of the subnetwork origin destination matrix error between the base and impact scenarios. Using the linear regression models from the Austin network provides reasonable recommendations, but there may be differences between the Austin and Dallas network that are not taken into consideration. An examination of the volume to capacity ratio for Jackson Street reveals that the base model predicted low volumes resulting in volume to capacity values below 0.1. It is possible that a greater network connectivity interacts with such a low volume to capacity ratio to prevent rerouting, or the capacity reduction could be less than the available capacity in the base scenario. This is indicated by the lack of correlation between rerouting at the subnetwork boundary and subnetwork size in each of the Jackson Street Scenarios. The smallest impact scenario for Pearl Street also has no correlation with subnetwork size.

For the larger impact scenarios, there appears to be a similar trend between subnetwork size and rerouting at the boundary as detected in the Austin network. As subnetwork size increases, rerouting at the boundary decreases. It was anticipated that the greater connectivity of the Dallas network would provide more alternatives for rerouting; thus, the linear regression model was expected to under predict the quantity of rerouting. This is demonstrated in Table 16 by the much larger true error than the predicted error. It also appears that characteristics of freeway scenarios are not fully represented in the linear regression model generated from the Austin data.

Table 16: Comparison of the Final Impact and Base Model for Dallas Scenarios

| Scenario | Subnetwork Size | True Impact | Base Model | Impact Model | Model Difference | Impact Model Percent Error |
|---|---|---|---|---|---|---|
| Pearl, 1 , 25 | 5 | 6.98 | 2.05 | 1.89 | -0.16 | 73% |
| | 7 | 9.03 | 1.82 | 1.95 | 0.13 | 78% |
| | 9 | 6.21 | 1.59 | 2.01 | 0.42 | 68% |
| Pearl, 2, 50 | 5 | 12.51 | 2.86 | 4.24 | 1.38 | 66% |
| | 7 | 8.59 | 2.63 | 3.50 | 0.87 | 59% |
| | 9 | 7.73 | 2.40 | 2.76 | 0.36 | 64% |
| Pearl, 3, 100 | 5 | 15.75 | 1.17 | 6.40 | 5.23 | 59% |
| | 7 | 9.38 | 0.94 | 4.06 | 3.12 | 57% |
| | 9 | 6.83 | 0.71 | 1.72 | 1.01 | 75% |
| Jackson, 1, 25 | 5 | 3.92 | 4.10 | 3.17 | -0.94 | 19% |
| | 7 | 6.21 | 3.87 | 3.23 | -0.65 | 48% |
| | 9 | 8.75 | 3.64 | 3.29 | -0.36 | 62% |
| Jackson, 2, 50 | 5 | 4.13 | 2.84 | 4.22 | 1.38 | 2% |
| | 7 | 5.92 | 2.61 | 3.48 | 0.87 | 41% |
| | 9 | 5.93 | 2.38 | 2.74 | 0.36 | 54% |
| Jackson, 3, 100 | 5 | 5.64 | 3.38 | 7.77 | 4.39 | 38% |
| | 7 | 5.90 | 3.15 | 5.43 | 2.28 | 8% |
| | 9 | 6.44 | 2.92 | 3.09 | 0.17 | 52% |
| Woodall Rodgers, 1, 25 | 5 | 12.47 | -0.74 | -0.56 | 0.18 | 104% |
| | 7 | 8.24 | -0.97 | -0.50 | 0.47 | 106% |
| | 9 | 6.25 | -1.20 | -0.44 | 0.76 | 107% |

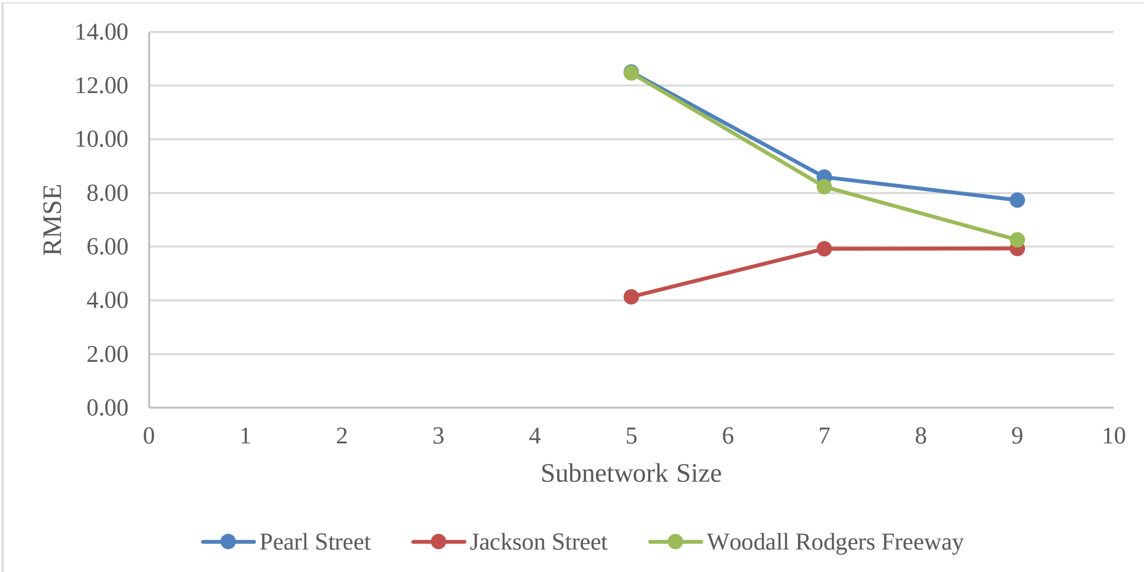



Figure 36: RMSE (Boundary Rerouting) versus Subnetwork Size, Dallas 2 Link Scenarios (1 Link for Woodall Rodgers)

Figure 36 indicates a decrease in rerouting at the subnetwork boundary as the subnetwork size increases for the mid-range scenario on Pearl Street and the Woodall Rodgers freeway scenario. The lack of volume on Jackson Street in the base scenario is likely the primary reason for such a low amount of rerouting at the extent of the subarea.

Table 17: Travel Time Validation of Dallas Subnetwork Recommendations, RMSE of Travel Times on Each Shared Link

| Scenario | Five | Seven | Nine |
|---|---|---|---|
| Pearl, 1 , 25 | 276.48 | 59.49 | 98.41 |
| Pearl, 2, 50 | 103.72 | 60.13 | 50.83 |
| Pearl, 3, 100 | 158.30 | 56.20 | 67.80 |
| Jackson, 1, 25 | 65.09 | 62.19 | 32.34 |
| Jackson, 2, 50 | 63.15 | 62.64 | 35.28 |
| Jackson, 3, 100 | 79.09 | 117.36 | 94.05 |
| Woodall Rodgers, 1, 25 | 112.15 | 60.38 | 49.94 |

Table 17 indicates that the lack of convergence for the Dallas network has a large impact on the ability of the subnetwork to produce accurate travel times. It appears that the Austin regression models may be used as a default method for determining a sufficient subnetwork size. However, calibration of local characteristics may be necessary when applying the methodology to different datasets. It is also possible that some metric may be able to account for the differences in the regions. Further examination of data from a DTA simulation with a lower level of convergence could provide information to calibrate such a metric.

A summary of all network characteristics is presented in Figure 37 through Figure 41. Each graph depicts a similar trend to the Austin network summary characteristics. The convergence does not appear to decrease as subnetwork size decreases for Pearl Street, which may be a result of the poor convergence of the downtown Dallas network.

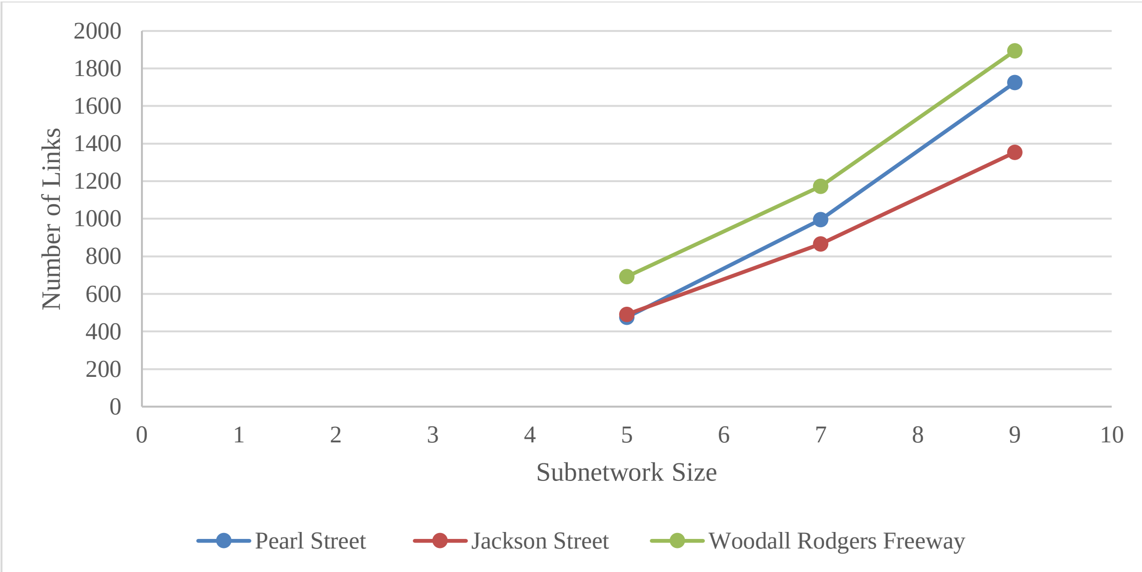



Figure 37: Summary of the Number of Links versus Subnetwork Size, Dallas

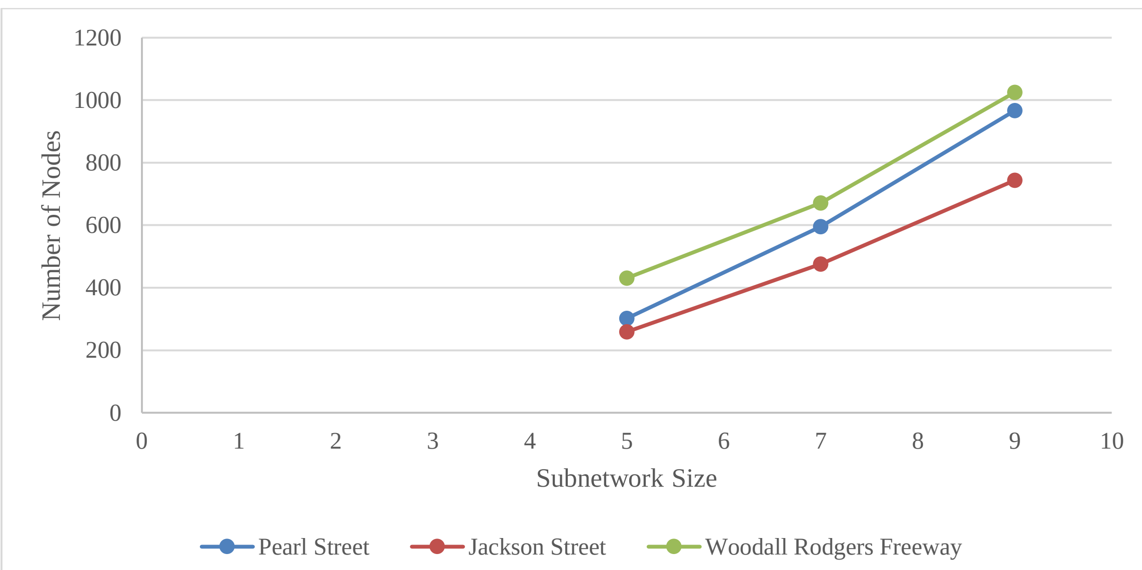



Figure 38: Summary of the Number of Nodes versus Subnetwork Size, Dallas

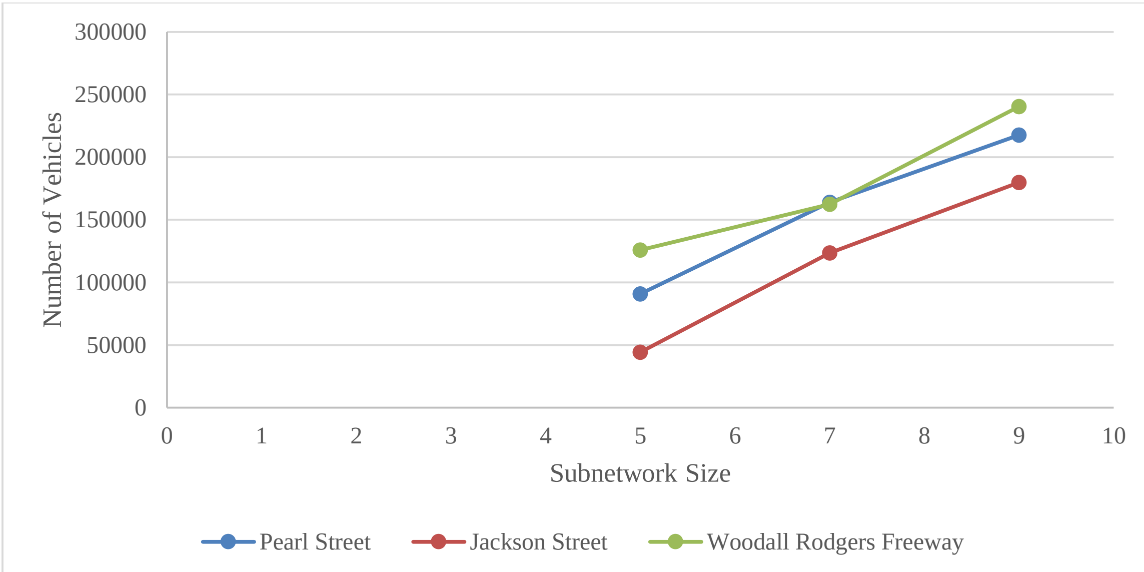



Figure 39: Summary of the Number of Vehicles versus Subnetwork Size, Dallas

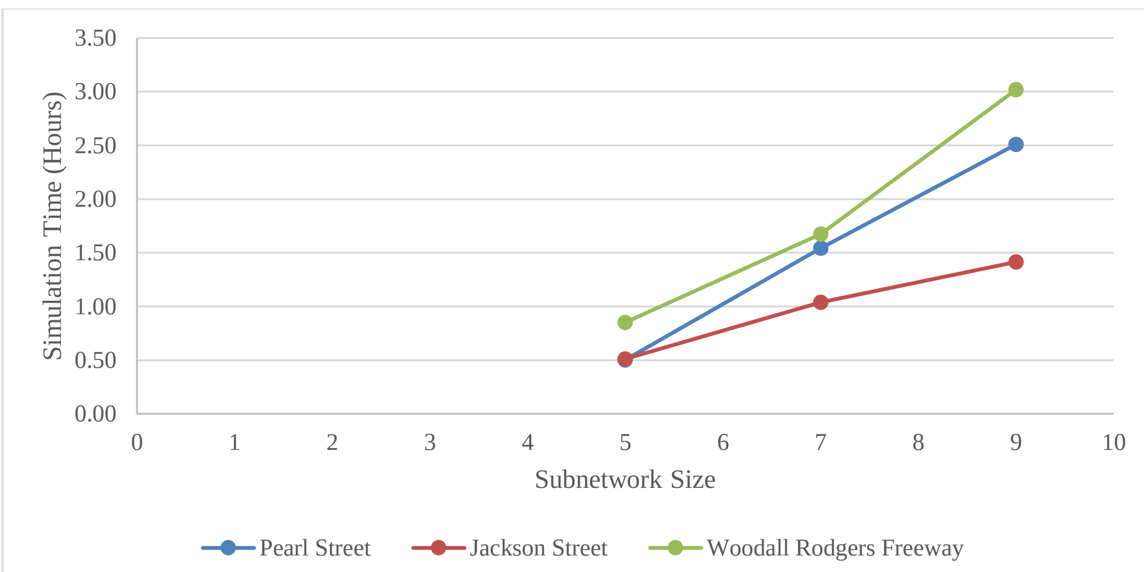



Figure 40: Summary of Simulation Time (Hours) versus Subnetwork Size, Dallas

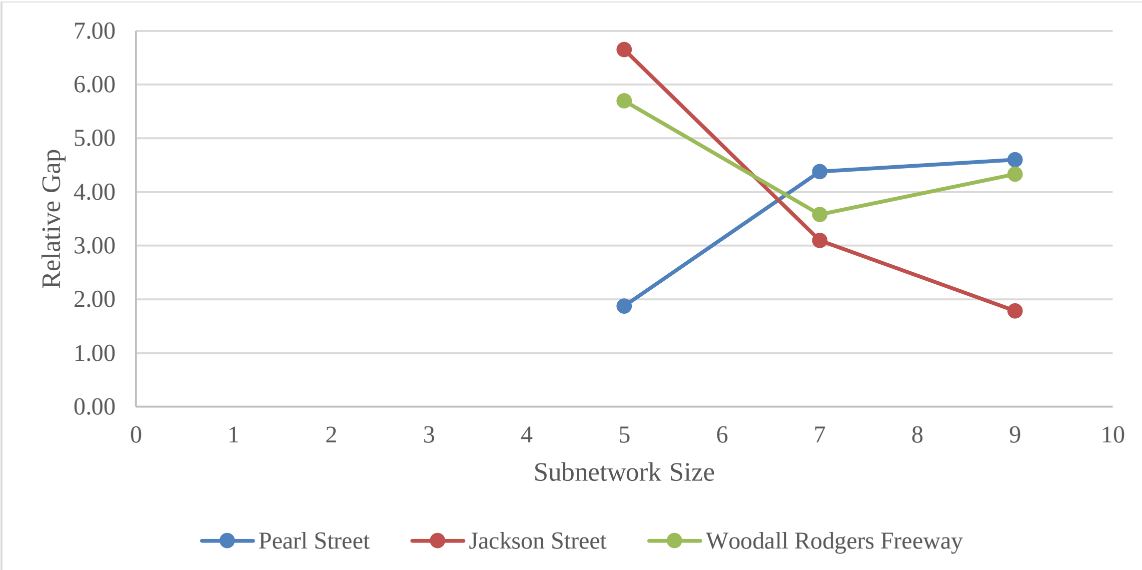



Figure 41: Summary of the Relative Gap versus Subnetwork Size, Dallas

Table 18: Representative Characteristics for Full Network versus Subnetwork, Dallas

| Spatial Extent | Number of Links | Number of Nodes | Number of Vehicles | Number of Cars | Simulation Time (Hours) | Relative Gap |
|---|---|---|---|---|---|---|
| DFW | 71721 | 31364 | 3119540 | 3015438 | 1,022.30 | 13.00 |
| Dallas | 16434 | 7408 | 858613 | 814312 | 72.20 | 7.15 |
| Downtown | 3958 | 1845 | 245752 | 230144 | 16.14 | 7.55 |
| Nine | 1725 | 967 | 217494 | 201651 | 2.51 | 4.60 |
| Seven | 996 | 595 | 163863 | 151114 | 1.54 | 4.38 |
| Five | 475 | 302 | 90782 | 84319 | 0.50 | 1.87 |

Table 19: Representative Percent of Full Network Characteristics for Subnetworks, Dallas

| Spatial Extent | Number of Links | Number of Nodes | Number of Vehicles | Number of Cars | Simulation Time (%) | Relative Gap |
|---|---|---|---|---|---|---|
| DFW | 100% | 100% | 100% | 100% | 100% | 100% |
| Dallas | 23% | 24% | 28% | 27% | 7% | 55% |
| Downtown | 6% | 6% | 8% | 8% | 2% | 58% |
| Nine | 2% | 3% | 7% | 7% | 0.25% | 35% |
| Seven | 1.4% | 2% | 5% | 5% | 0.15% | 34% |
| Five | 0.7% | 1.0% | 3% | 3% | 0.05% | 14% |

Table 18 and Table 19 reveal how great a difference using subnetwork analysis can make when dealing with a regional network as large as Dallas. The simulation time statistic is particularly revealing as the full network required over 1,000 hours while the order five subnetwork required only 30 minutes.

### 5.3 Practical Recommendations to Account for Nonuniformity

The methodology for subnetwork creation aims to capture the extent of congestion due to a network alteration. For dynamic traffic assignment, this means determining the quantity of rerouting that occurs with respect to time and space. This can be difficult due to the complex algorithms that are involved in the DTA process – this was emphasized with the analogy of water displacement in a channelized system from Chapter 3.1.

Nonuniformity in network topology introduces another level of complexity to generating a universal subarea selection technique. The uniqueness of regional traffic networks is not simple to classify because network structure characteristics can vary greatly. These physical network factors can include spatial resolution, connectivity, density, and geometry. Besides network topology other differences between regions can include population, demand, travel behavior, multimodal options, and traffic operations. These unique features can create noise in the results.

Homogeneity in the roadway network may be an effective characteristic to remove noise. Tsekeris and Geroliminis (2014) examine network structure effects on traffic congestion from an aggregate, economic perspective using a uniform grid network. This type of experimental control can frame the concept of a network connectivity metric as a deviation from complete homogeneity. A proper network topology metric should decrease as the spacing of a uniform roadway grid increases (the network becomes more sparse) or as the network form deviates from a completely connected grid.

Qualitatively, the ideal scenario is a perfectly uniform grid network. That means all links of the uniform grid network have equal lengths, number of lanes, capacities, and speeds. This would be the equivalent scenario to the rock being dropped in the middle of a uniform puddle (from the analogy in Chapter 3.1). In this case, it is expected that the impacts should spread symmetrically about the location of capacity reduction. As soon as the network deviates from the ideal grid, these effects become more complex and tracking them becomes more difficult. Therefore, the user should be aware that analyzing portions of the network that are nonhomogeneous may reduce the accuracy of this subnetwork characterization methodology. However, this does not necessarily prevent this method from being appropriately applied.

There may be ways to quantitatively account for the differences in network structure. Classification of transportation networks for traffic analysis has been limited to primarily qualitative measures. The FHWA Traffic Analysis Toolbox Volume 2 (Jeannotte et al., 2004) provides a nominal classification of study area, or geographic scope, limited to: isolated location, segment, corridor/small network, and region. These general terms add no valuable distinction between regional differences. Development of network topology characterization has been addressed, foremost, by the transportation planning community in pursuit of urban form metrics. These metrics tend to focus on connectivity from a level of service perspective such as mobility or accessibility. Since these metrics do not address the fundamental principles of traffic flow they have little capability in providing a reliable quantitative metric.

Zhang and Kukadia (2005) provide a grouping of commonly applied metrics for urban connectivity including: geometric, context sensitive, and behavioral measures. The geometric measures attempt to address connectivity by determining the ratio of the number of links to the number of nodes in the network.  A better geometric measure is the use of roadway density, miles of road per square miles of area. Context sensitive measures hope to capture the population characteristics of surrounding area, and since this is already accounted for in the origin-destination demand input has less value to characterizing topology for traffic theory. The behavioral metric is defined as accessibility in the form of a gravity model relating origin-destination impedances to opportunities in the traffic analysis zones. Dynamic traffic assignment aims to more accurately define these network impedances and uses destination demand to determine the desirability of zones, which makes accessibility a less valuable measure in this context.

Some additions have been made to Zhang's classification of network structure. Jenelius (2009) identified many of the same network structure measures as Zhang, but labeled the geometric measures as "link redundancy" (links to nodes ratio). Jenelius also mentions a potential metric of network scale, the average link length. The average link length may be biased, or influenced, by the arbitrary method used to code the individual links in the network, but is an important consideration for spatial analysis of traffic networks – and could have implications for the connected order method. Jenelius discusses the use of roadway density (miles of road per square mile), and suggests that higher density should increase the availability of the number of routes. Lowry and Lowry (2013) identified a more thorough list of urban from metrics including density (population), centrality (gravity model), accessibility (street connectivity), and neighborhood mix (demographic diversity). The four metrics they investigate for street connectivity are the ratio of links to nodes, mean neighborhood block perimeter, ratio of cul-de-sac to streets, and median length of cul-de-sac. Some improvement is needed for these planning metrics to make them more applicable to engineering problems.

Road density may be the best characterization of urban form from a traffic engineering perspective, but is still limited in the capability to represent the fundamental traffic flow parameters. In a general sense, measuring the total length of roadway in a subnetwork captures an aggregate measure of the denominator used to calculate traffic density (vehicles per unit length). Another fundamental traffic parameter, flow, could be partially captured if capacity was included in the roadway density calculation. This technique has the potential to differentiate roadway classes by including the number of lanes and the differing capacity per lane based on HCM procedures. For example, typical traffic engineering simulation has a much higher capacity coded for freeway links compared to arterials. A metric including these characteristics may take the form of lane miles per square mile, which will be referred to as road lane density, or lane mile capacity per square mile, which will be referred to as road capacity density.

### 5.4 Addressing Nonuniformity in the Dallas and Austin Regional Areas

The network data used for this analysis includes two different regional networks with differing types of connectivity, the Dallas and Austin regional areas. Road lane density and road capacity density were calculated for each of the subnetworks created (in both regions). The areas used for the subnetwork were defined by the minimum bounding geometry. The most appropriate geometry method to use is the convex hull, seen in Figure 42 since it uses a polygon to connect the minimum area containing the furthest extents of the network links. The lane miles term can be calculated by summing the product of each link's length and number of lanes. The lane miles capacity term can be calculated by summing the product of each link's length, number of lanes, and lane capacity. Then, the area of the convex hull of subnetwork links can be used to normalize the quantity metrics for the roadways.

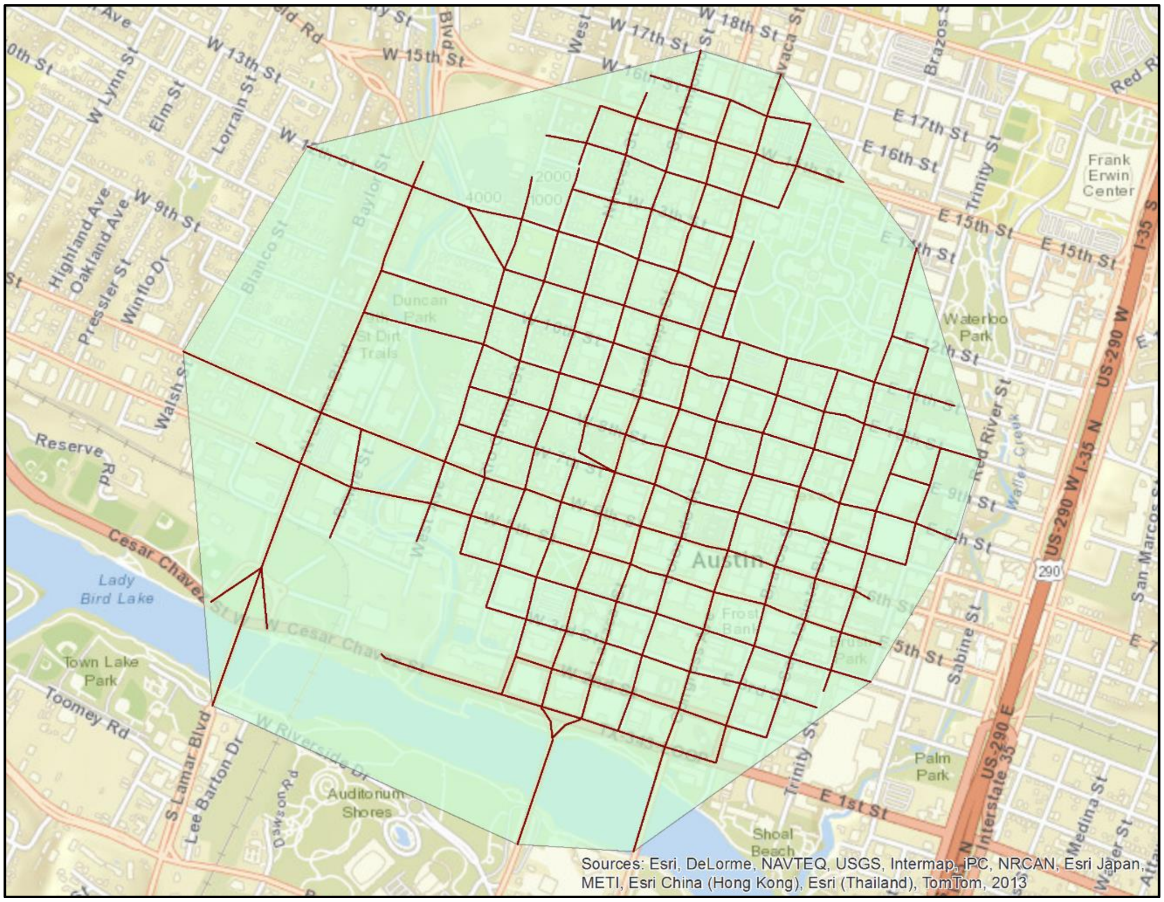



Figure 42: Convex Hull Minimum Bounding Geometry for the Calculation of Area used in Network Connectivity Metrics

The aim of this analysis is to determine the nature of the relationship between the connectivity of the subnetwork to the quantity of rerouting that occurs at the subarea boundary, which has been defined as the root mean square error of the induced boundary demand. For each subnetwork size and number of links impacted, each impact location was plotted with the connectivity metric on the x-axis and the measure of rerouting on the y-axis. Due to the differences in the level of convergence for the two regional networks, they will be displayed separately to identify potential trends. Other factors that are included in the linear regression equations are no longer being controlled, leading to a great deal of variation around the best fit line. However, Figure 43 through Figure 46 indicate a likely influence of the traffic network connectivity on rerouting.

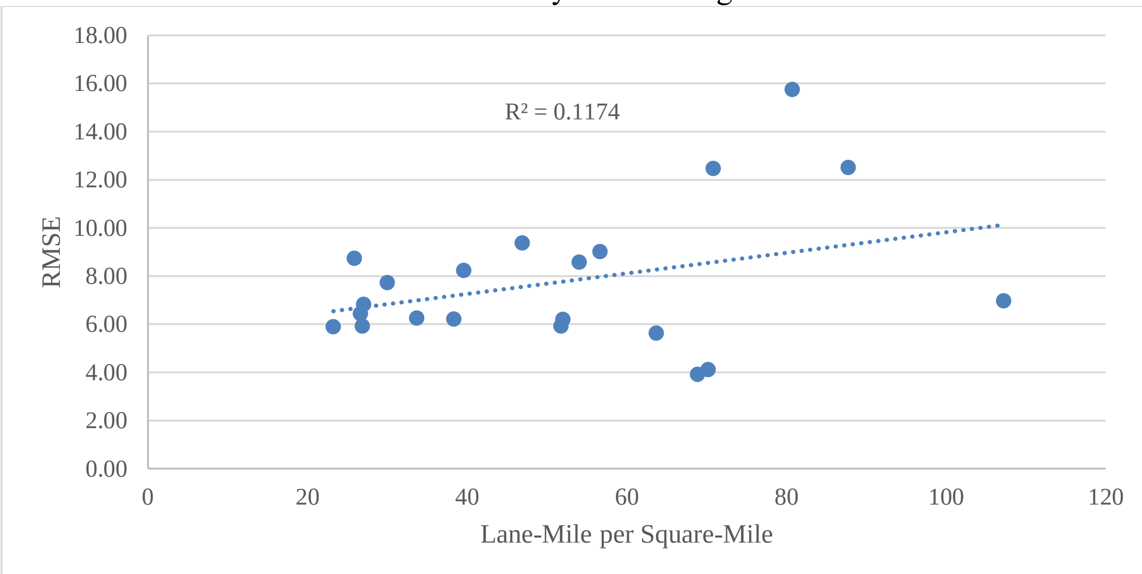



Figure 43: Relationship of Network Connectivity to Vehicular Rerouting at the Boundary of the Subnetwork, Road Lane Density for Dallas

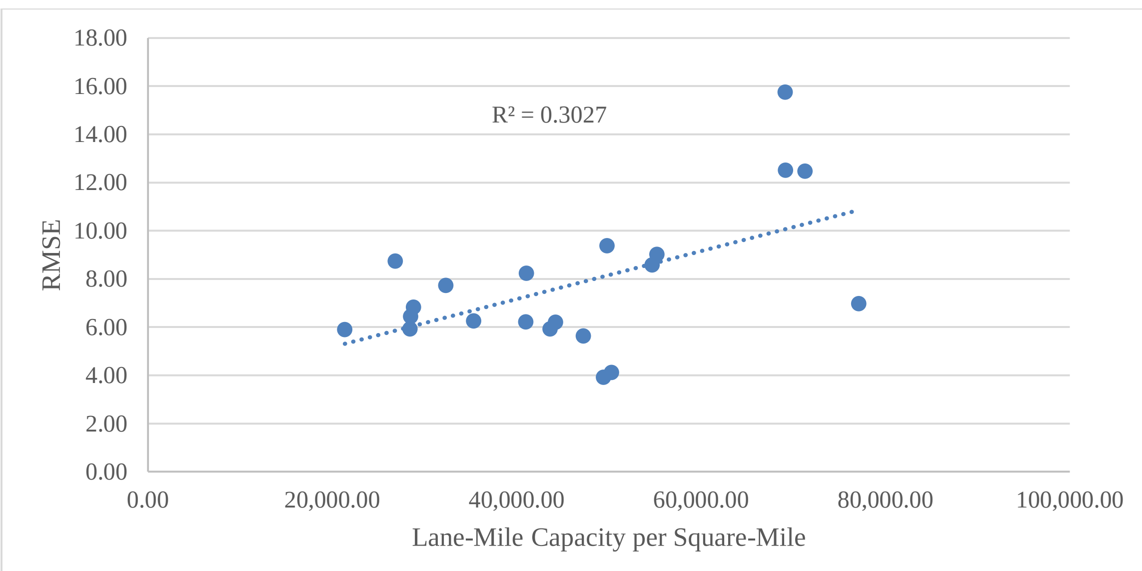



Figure 44: Relationship of Network Connectivity to Vehicular Rerouting at the Boundary of the Subnetwork, Road Capacity Density for Dallas

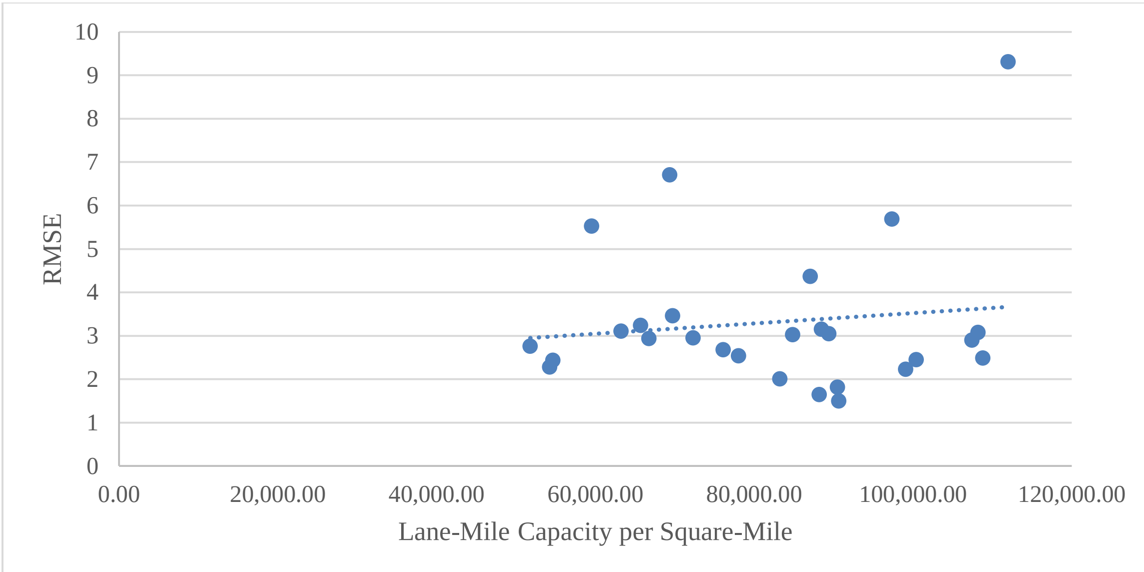



Figure 45: Relationship of Network Connectivity to Vehicular Rerouting at the Boundary of the Subnetwork, Road Lane Density for Austin

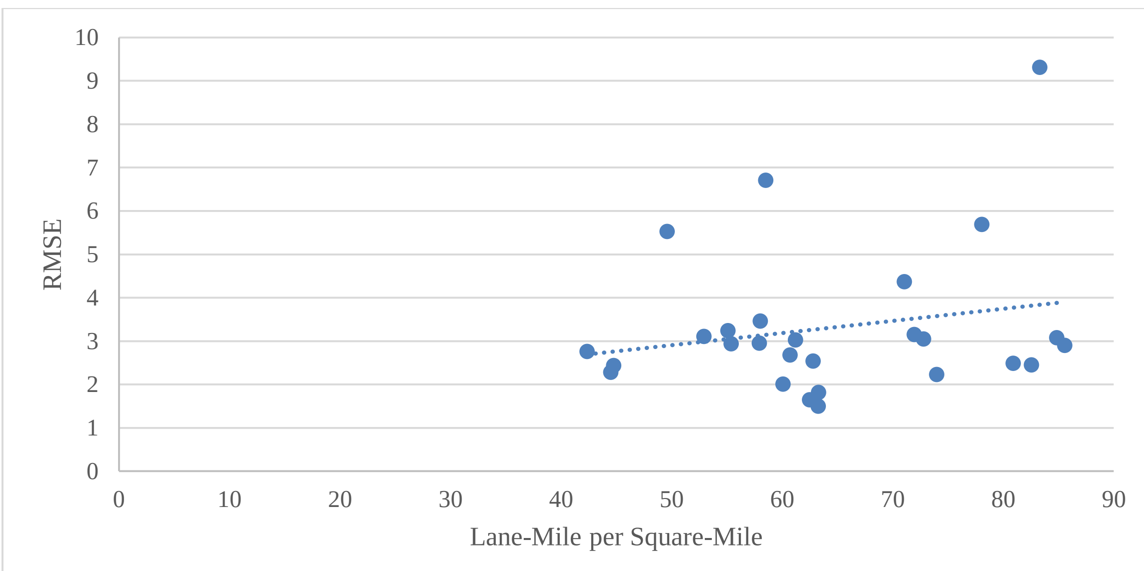



Figure 46: Relationship of Network Connectivity to Vehicular Rerouting at the Boundary of the Subnetwork, Road Capacity Density for Austin

The general trend reveals that rerouting at the boundary increases as network connectivity increases, which follows intuition and the assertion by Jenelius. Figure 43 and

Figure 44 indicate that including capacity in the calculation of network connectivity increases the predictor capability of the metric, measured by the higher coefficient of determination (0.30 compared to 0.12). There is a great deal of variation in each Figure 43 through Figure 46, particularly Figure 45 and Figure 46. This is due to the many other factors that characterize each scenario, particularly the impact scenario magnitude and level of congestion in the subnetwork. However, these factors are included in the linear regression model and including this network connectivity metric could improve the model as well as address the variation seen in Figure 43 through Figure 46.

A further examination of this analysis is recommended for future research as network connectivity metrics for traffic simulation could potentially improve network analysis. Such a metric can help to estimate rerouting, differentiate between network topology, and recommend/measure spatial resolutions for subareas of analysis. Further attempts to include the traffic network connectivity metric into the predictive linear regression model would require more data from a greater number of different regions.

The use of the convex hull also has potential to enhance the connected order selection method. An additional step to the connected order selection technique of generating the convex hull area for the subarea links, then adding links within the convex hull will prevent discontinuities in the subnetwork. In Figure 42, it can be seen that the connected order has the potential to leave gaps in the network, which could be corrected with the use of the minimum bounding geometry geospatial tool.

Geographic bias is a major problem for any spatial statistical analysis. Since transportation analysis is inherently spatial, accounting for this bias in the engineering and planning simulation process can increase the reliability of model results. The primary consideration for geographic information analysis is the modifiable areal unit problem

(O'Sullivan, 2010). Most geographic data is aggregated when analyzed, and this was especially true for this study since each measure was calculated with respect to the data within the subarea. The aggregation of this data at different scales can cause potential issues with regard to (1) defining the appropriate aggregation scale and (2) the impacts of using different aggregation scales. This study addresses both of the issues through (1) the use of an unbiased geographic area method of increasing subnetwork size and (2) proposing the use of a network connectivity metric that can account for the difference between aggregation scales.

With any spatial metric there is an inherent geographic bias associated with the data. However, the controls in this experimental setup have limited the issues associated with geographic data. Other concerns still exist with the input data for this analysis, including Jenelius' mention of the average link length in the network. The biases that can occur within the initial data creation step are a critical reason why there is a great need for standardizing geospatial datasets used for transportation analysis. The development of databases for traffic analysis zones, centroid connector placement, link location and length, link capacities, and other network and demand characteristics must be controlled to reduce bias in transportation modeling and analysis.

### 5.5 Summary

The Austin model has been validated through a comparison of the travel time prediction capability of the different size of subnetworks. The ability of the linear regression model to predict a sufficient subnetwork is supported by the travel times being fairly similar at the recommended subnetwork size. The predicted sufficient subnetwork size from the linear regression also matches the results of the comparison test.

A Dallas DTA model was simulated with scenarios comparable to the Austin case study used to develop the subnetwork methodology. There may exist network characteristics that are not captured by the regression model built for Austin. However, there is still potential for using the linear regression model in its current state to predict an appropriate subnetwork size. If users wish, they may adapt the model to their region and calibrate the coefficients on the parameters.

It is possible that another metric, which describes network topology, may be used as an explanatory variable. This metric could account for the differences seen in the different network regions. If a metric that characterizes the network could be identified, it would prevent the data collection effort required by the comparison linear regression methodology. There are several network structure metrics in literature, but they rarely address traffic operations by incorporating the fundamental flow characteristics. Road lane density and road capacity density are introduced here to capture network structure with respect to traffic flow principles.

# Chapter 6: GIS Implementation

Geographic information systems have changed the way engineers store data, interface models, and perform analysis. Transportation was in great need of this new technology because topology dictates traffic operational interactions. As high quality geospatial data becomes a standard in the industry, considerations should be made so that all disciplines solving transportation system problems can coordinate their efforts. Standardizing a geodatabase format will help to realize the full potential for this data to improve modeling and analysis procedures. Spatial analysis and interfacing different modeling software inputs and outputs are only a first step in utilizing geographically referenced data. GIS can be used to automate routine processes and provides potential for interfacing with new real time ITS data. A tool for interfacing dynamic traffic assignment outputs and traffic control plans is described in this section to demonstrate the capability of GIS and considerations for the future. These applications are only a small component of the transportation field transitioning to digital solutions, rather than paper manuals. Oversight will be necessary to ensure decisions based on software outputs are objective.

## 6.1 Geospatial Data Considerations for Traffic Analysis

In an attempt to capture more factors of reality the complexity of theoretical models have increased. Simultaneously, as the tools change to describe ground truth more accurately, so has the level of detail for our data and the software used for processing. GIS capabilities have grown as a result of needs from a variety of different fields. For example, coordinate system science provides a means to standardize representation of geographic location and ensure the accuracy of geospatial calculations. The adaptation of the World Geodetic System (WGS 1984) for global positioning systems (GPS) created a new standard; although, many public GIS datasets rely on the North American Datum of 1983

(NAD 1983). Awareness of these ellipsoidal referencing systems and the corresponding Cartesian projections is critical in the development of a centralized database. Most GIS packages allow for translating datasets between coordinates to provide consistency for georeferencing and geocoding. Conveniently for traffic engineers, geocoding is commonly based on roadway networks. Using the spatial relationships of data features allows for automation of operations on their attributes.

A consistent framework for geolocating data features is the backbone of spatial analysis. With an accurate representation of the topology, transportation professionals can apply geostatistics that have already been developed for other fields including geography and environmental sciences. These tools provide insight on the street network and its relationship to other relevant geodata including pavement management information systems (PMIS), land use, crash records, and traffic analysis zones (TAZ). The real power in this platform is the ability to customize automated procedures consisting of preexisting tools and user defined subroutines. Laborious processes can be simplified with a script, saving the user time that could otherwise be spent improving data analysis.

Specifications for data structures, quality, and formats could help to provide a level playing field for traffic analysis. Traffic engineers could rely on data sharing formats similar to the general transit feed specifications (GTFS) to ensure interoperability. It will be important to orient the Highway Capacity Manual (HCM) analysis and other FHWA procedures to focus on topographical interactions of transportation systems. Publication of future professional manuals should account for the digital solutions needed to address the new form of traffic data. Intelligent transportation systems will progressively be adapted as the primary source for data and the procedures outlined in engineering guidebooks should reflect these digital formats.

Major limitations in the available datasets can be addressed through the TFTN initiative, FHWA, MPOs, state departments of transportation (DOT), and local agencies. The question of the proper data structure is not a trivial one. Accounting for the needs of multiple disciplines makes a universal database format difficult to establish, and issues with the absence of desirable data adds a level of complexity. Two major issues that a standard GIS dataset should address for transportation are the attributes required for different resolutions of traffic simulation and a method for integrating traffic signal timing plans into the network database. Once an inventory of the transportation system have been addressed, then the advantages of applying data from the digitally connected infrastructure data sources can be realized. Figure 47 is one example of how data integrity is rapidly improving.

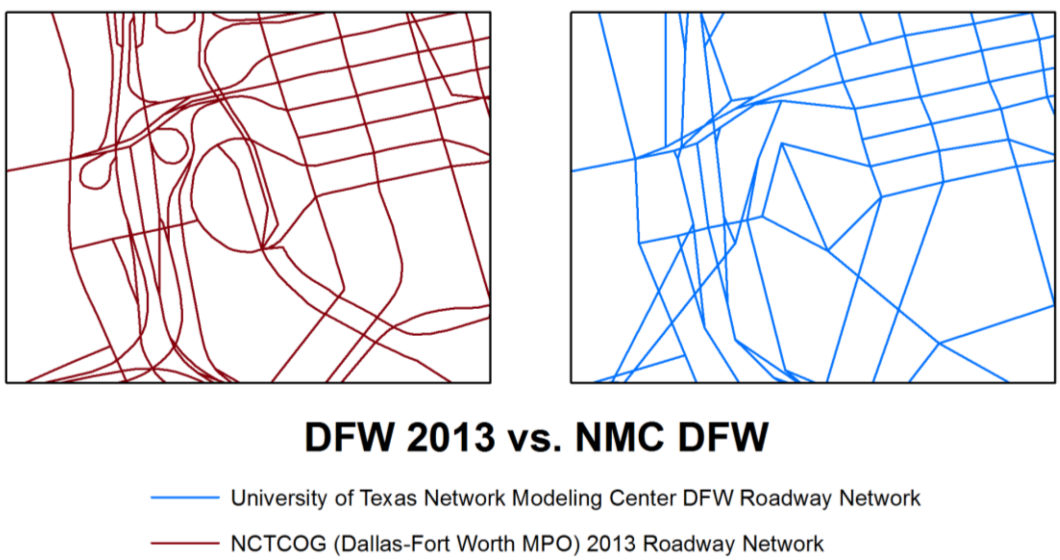



Figure 47: Difference between New and Old MPO Networks

## 6.2 GIS Considerations of the Subnetwork Procedure

GIS has a large role in organizing, editing and analyzing traffic simulation results. As the models used within the simulators become more complex, visualizing changes in network performance becomes even more important. Better quality data will help to improve the accuracy of these advanced traffic models. Throughout this process it is valuable for traffic engineers to understand the fundamentals of GIS and for the GIS developer to understand the needs of traffic engineers. The commonality of the GIS data format and the development platform make it a useful tool for customizing geospatial procedures.

In order to take advantage of the power of DTA, visualizing the routes that are used by vehicles can provide unique insight into the output of the simulation. This can be done in several ways. The more common way is to investigate all routes used between an origin and a given destination. This method can identify the most used route between a particular origin-destination (OD) pair and what alternative routes are available for the vehicles using that OD pair. For a traffic control analysis it can also be useful to select the link(s) that are going to be modified and find the set of paths that use that link. In this case it is easier to analyze the results if the number of routes is limited to the top ten route volumes. Then, the vehicles that use these routes can be tracked in the impact scenario to determine the alternate routes they use. This can reveal the network distance from the modified link where the impact has spread. This was one of the initial methods mentioned in Chapter 3, and a map representing the distance calculation for this technique is represented in Figure 48. Even though it was not implemented for the subarea selection, it is still a useful tool for investigating impacts.

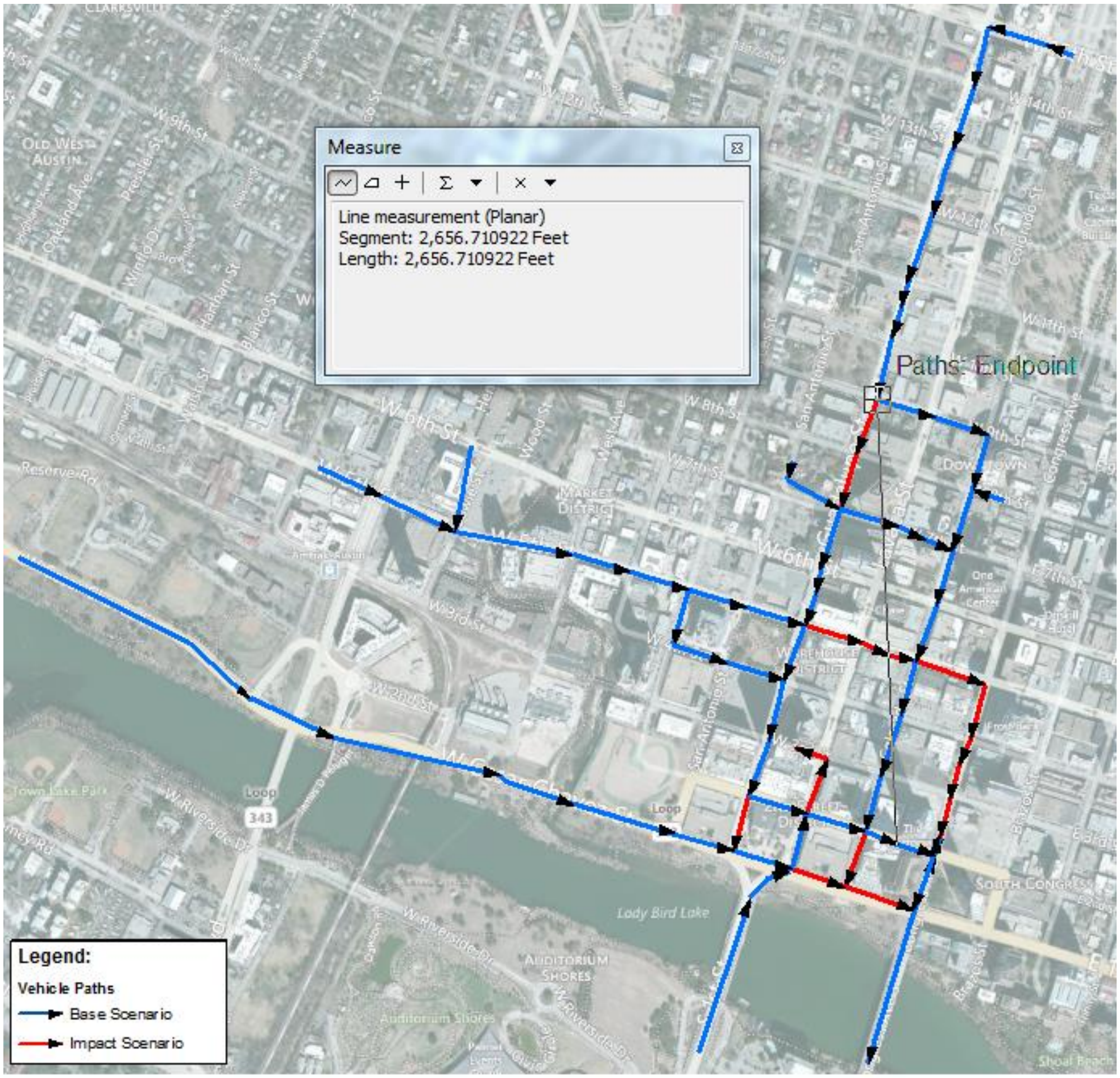



Figure 48: Visualization of Paths Rerouting Vehicles around the Impacted Link

The investigation of absolute distance for the furthest rerouting directed the method in a new direction. In geographic analysis terms, the focus transferred from straight line, Euclidean distance (as the crow flies) to network distance. Network distance uses links as a measure of length. The GIS components of the subnetwork procedure introduce in this dissertation use spatial analysis tools and, assuming that the DTA model data is properly formatted, the GIS algorithm could be run on any transportation network. The subarea selection procedure referred to as connected order may be automated by comparing the geographic coordinates of the nodes comprising the ends of the links. A user may select a link, then the nodes connected to this link are automatically selected using a select by location tool. Then, the connectors and links emanating from the selected nodes will be added to the selected features. Finally, the centroids that are used by the selected connectors will be added. This procedure will automatically iterate until the appropriate number of links extending (user input) from the primary selection are included. This standardized method of creating a subnetwork was designed to enable investigation of how far the impacts extend, and provide a means for automating this process. Figure 49 and Figure 50 are the GIS interface for the connected order tool.

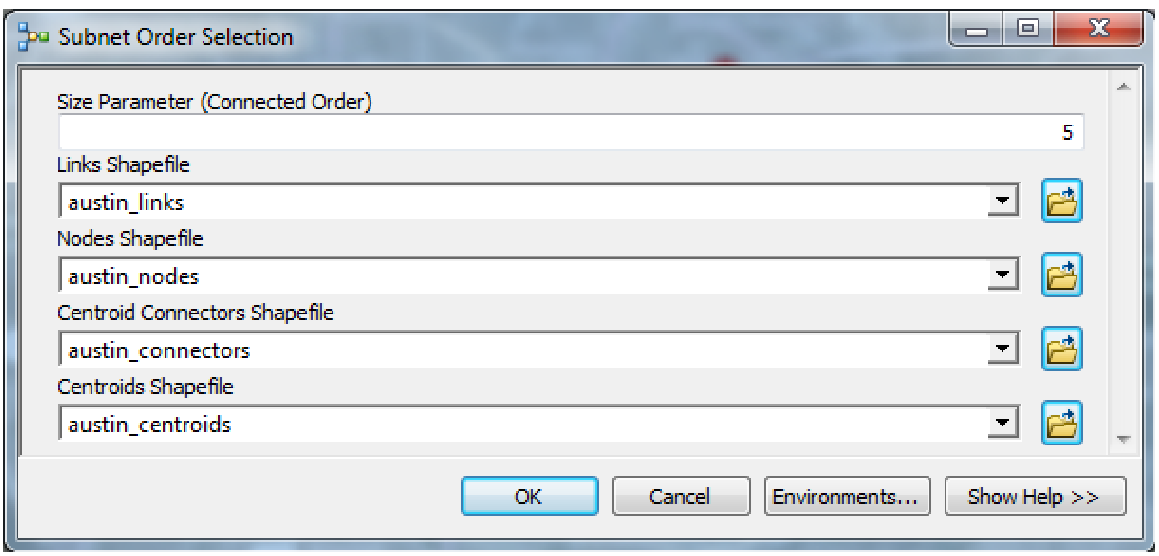



Figure 49: User Input Prompt for the GIS Connected Order Selection Model

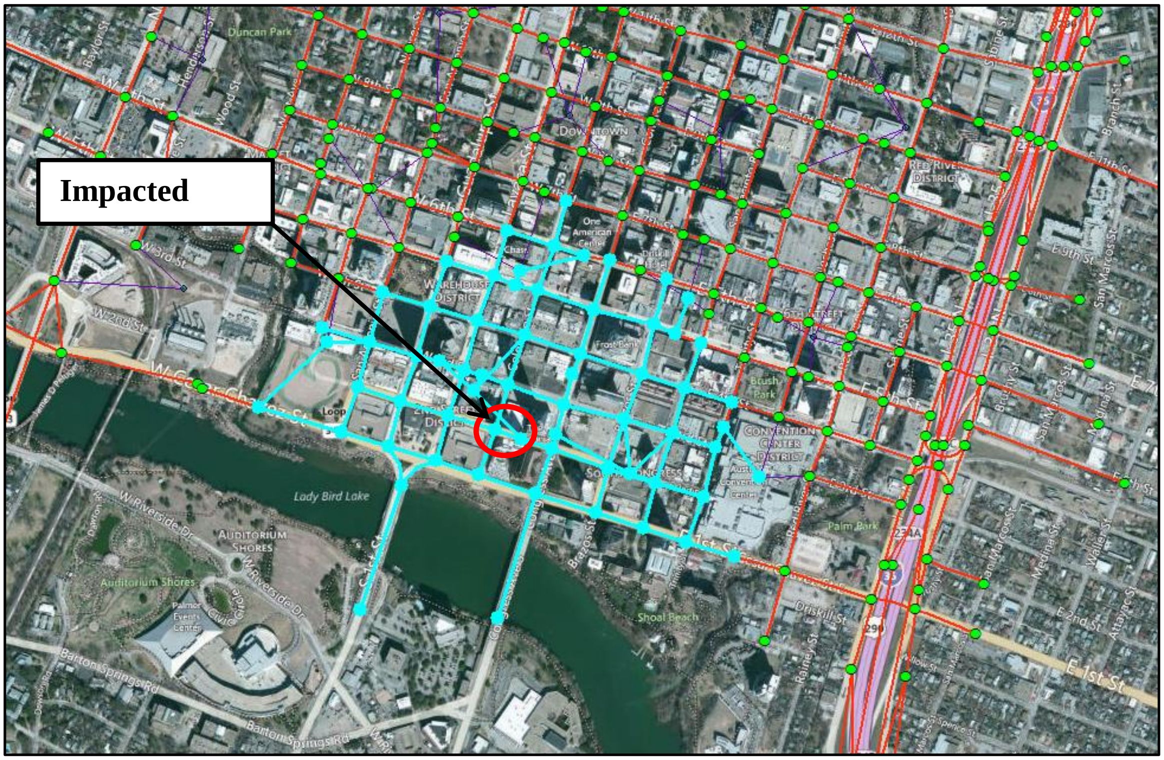



Figure 50: Sample Connected Order Selection with Size Parameter of Five in GIS

### 6.3 Creating a GIS Tool for Utilizing DTA Subnetwork Data

The first step in creating a customized tool is to synthesize the procedure and identify inputs, processes, and outputs. The example flowchart in Figure 51, provides an organizational overview to the scripts that were developed for a preexisting traffic control planning tool (Bringardner, 2012). Each major step may require several actions, and each action may be carried out with a tool set provided by the development platform or may be automated in GIS. Online forums are the best resource for code templates and advice for GIS development. As with all programming, debugging errors will be the final step to this process. Troubleshooting is even more prevalent with geodatabases due to the quantity of manually input data and the high probability of human error. These data issues cannot be overlooked in the development stages of a new algorithm and can be minimized if the proper data sources are found.

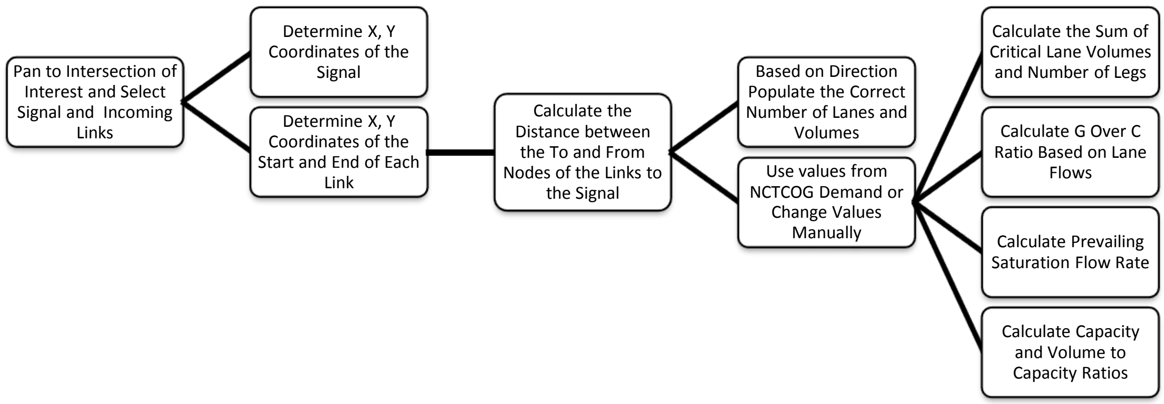



Figure 51: Sample Flow Diagram for the Example GIS Traffic Engineering Tool

In light of the TFTN (Transportation for the Nation) planned timeline for completing a consolidated centerline and transportation GIS database, choosing the best input data for analysis often relies on contacting the regional metropolitan planning organization (MPO). MPOs are often one of the few organizations that update network GIS data with attributes necessary for traffic modeling. Consultants and contractors should begin to digitally upload the data used for their projects to streamline the quality control of the TFTN, or other centralized datasets. Requiring their updates for various projects as a digital deliverable could help remove redundancy and keep track of previous decisions for future projects – much like a set of plans documents engineering decisions.

Interacting with the data requires an understanding of the attributes associated with the features of interest. Transportation network models are made up of links and nodes, represented in GIS by lines and points. Geographic coordinates are assigned to ends of the roadway links and each intersection node. Most traffic models require a directed network, or separate representation for opposing flows of traffic on a roadway. Desirable attributes for the links are the name, unique identifier, number of lanes, lane configuration, speed limits, geometry, traffic volumes, and vehicle type distribution. The level of detail and the method of labeling these attributes varies greatly between datasets. Node data should classify name, unique identifier, the type of stop control, contain timing plans for signalized intersections, and coordinate turning movements of the intersection links.

### 6.4 A DTA Integrated GIS Tool For Traffic Control Plans

An example development of a GIS tool is described in this section to present an approach for integrating transportation models and geographic information software. This tool allows for a user to make sketch planning calculations for capacity changes and interface with a dynamic traffic assignment (DTA) model to predict routing in the case of a traffic control plan. If a lane closure is imposed on the network, then vehicle routing is likely to change near the capacity reduction. To investigate these impacts the GIS model incorporates HCM procedures for calculating roadway capacity. Functions have been designed to import and export outputs to and from DTA software for a more accurate representation of the after scenario network flows.

The automation of these calculations, which would otherwise require manual inputs in traffic software, relies on the spatial relationship of the road network. Identifying the signal location and incoming links through nearest neighbor calculations simplifies the task of extracting the relevant information from the database. To allow for the interoperability of new input datasets, in particular DTA updates, all calculations are based on the unique identifiers of the features. The inherent structure of the database format requires an association between every feature and each attribute in the dataset. Each process operates on the features attributes connected to their unique identifiers to allow for reproduction, flexibility, and improvements.

The integration of the DTA subnetwork procedure is essential to the GIS tool for improved network alteration analysis. A user could select a link(s) that is under study for potential capacity alterations. The user could then input the range of capacity reduction impacts that they wish to implement on the work zone roadway. Using this information and output data from a validated full regional model run (volume to capacity ratio), the predictive linear regression algorithm can determine the appropriate subnetwork size. It does this by automatically extracting data and calculating the estimated rerouting associated with possible network sizes. If the user wishes, they could use the default recommendation or they could choose an option to view the estimated errors associated with subnetwork size. The automated system could then produce the final input data required to build and run the subnetwork in the DTA software.

In the future, if a coefficient is calibrated for the network connectivity metric in the linear regression equation this could also be automated in the recommended subnetwork size specification. A GIS model has already been constructed for calculating the area using the minimum bounding geometry and the corresponding road lane density and road capacity density. These automated procedures can send the calculated network connectivity parameter to the linear regression model for creating subnetwork recommendations. This would prevent the user from having to input any information other than the scenarios they are interested in analyzing.

The strength of the database calculations is the ability to generate digital results that can be used by other software. In this case, initial input volumes for the previously developed traffic control planning tool came from a static traffic assignment model, which means calculations were based on an average hourly volume per link. Figure 51 reveals the GIS processes that were used to calculate volume to capacity ratios for an intersection involved in a network alteration. Although the capacities were recalculated to adjust for the network changes, the volume is static and does not represent the corresponding changes in traffic flow. Combining the dynamic traffic assignment model results allows for more detailed analysis and data that reflects the new state of the network. The GIS processes include selecting the subnetwork and calculating the volume to capacity ratios on desired links with the DTA output data reflecting rerouting. The subnetwork procedures allows for short simulation times and the generated output from the DTA model can be used for the subarea in the traffic control planning tool.

***Procedure for Using DTA Subnetwork Results for Traffic Control Planning***

**STEP 1:** Select and extract subnetwork data elements using the GIS model interface

**STEP 2:** Simulate the subnetwork in a DTA software platform

**STEP 3:** Extract updated volume results from the DTA model

**STEP 4:** Import DTA output for the subarea into the traffic control planning tool

**STEP 5:** Use the GIS traffic control planning tool to calculate volume to capacity ratios

### 6.5 SUMMARY

The process of creating a GIS traffic control planning tool was described as a generalized method that may be used to implement other GIS decision making tools for traffic simulation data. Each component of this subnetwork methodology was automated within a GIS software platform. The issues associated with integrating traffic simulation and GIS tools were identified. The most important concerns for GIS data are database structure, interoperability of functions, and geographic relationships. Geospatial data considerations are becoming more important as engineers rely on large datasets with minor details that have a large impact on simulation and modeling results. Developing a standard, digital structure for this data is the primary means for improving the integrity of transportation analysis.

# Chapter 7: Conclusion

The aim of this dissertation was to add another level to the current understanding of spatial interactions in advanced traffic models. Geographic resolution analysis is particularly complex because of the potential for spatial interaction to bias typical analysis. Dynamic traffic assignment is at the forefront of traffic engineering simulation tools, and contains a great deal of spatial complexity. DTA also operates within the time domain, which means changes must be tracked over time. A robust set of tools were needed to account for the model variation in time and space.

## 7.1 Summary of Contribution

This methodology created a means to identify a sufficient subnetwork, which means that it is large enough to accurately simulate traffic and small enough for significant reductions in computer time required to process the algorithms. The key to this technique is statistically comparing the subnetwork demand inputs. If the subarea is run in the same DTA platform, then the same demand matrix input should generate the same output. After identifying the trends involved in the growth of subnetwork size, a more sophisticated method was developed for relating the inputs in the full network to the outputs used by the full network. This contribution extends beyond subarea analysis because it identifies factors that influence the propagation of congestion through traffic networks. For instance, this process identified that the current level of congestion in the base scenario, represented by the volume to capacity ratio, is a good indicator of the quantity of rerouting caused by new congestion as a result of network alterations. It was also discovered through simulation data that network topology likely has an influence on the possibility of rerouting. A metric that quantifies the network structure and connectivity could be applied to this subnetwork analysis, but can also help to compare and quantify differences between regions.

Accurately identifying a sufficient subnetwork is critical not only for practical implementation, but can also help develop dynamic traffic assignment and traffic simulation research. The use of this method will allow practitioners to understand the error associated with the selected subnetworks for traffic analysis. It may also be used as an automated tool to more quickly create subnetworks for multiple scenarios. The use of the subarea selection techniques proposed here may also help the implementation of experimental algorithms. Typically, researchers rely on abstract networks, like Sioux Falls, to test new ideas. With assurance of the quality of subnetwork data supplied by this dissertation, researches can use real world scenarios in areas they are familiar with to test the particular scenarios they are attempting to address. The results of the cause and effect regression analysis has also identified issues that must be considered when examining multi-resolution analysis. In particular, network topology and geographic study area are critical issues in transportation network analysis.

### 7.2 Summary of Comparison Technique

Generating the methodology for subnetwork analysis required identifying the appropriate techniques for spatial and temporal investigation. A review of typical tools in transportation systems analysis found shortcomings for particular methods when dealing with varying geospatial areas. For instance, these windowing techniques create an issue for link level statistics because simulation at a smaller spatial scale generates additional variation in results that already exists from the demand input extracted from the full network. Another example is the limited capability of a utility base logit model for route choice. Although it can be effectively used in network traffic assignment, when applying it to a subarea it requires the generation of more data outside the window of analysis. This means that the destination choice logit model is a less effective means of improving subarea analysis because the data required for its input is comparable to simulating the DTA process on a larger area. These shortcomings were identified and guided the research in a new direction for adequately addressing the needs of subareas.

The primary contribution of this method is a novel application of a metric for identifying the spatial extent of impacts in DTA. Rather than investigating the outputs of the subnetwork for identifying what is sufficient, the methodology focused on the inputs. Most commonly in practice, the input to the subarea is the induced boundary demand extracted from the vehicle trajectories of the full network. This transition in focus from link based statistics was a significant step in the subarea methodology. This study found that the standard measure of variation, root mean square error, was the most appropriate for subnetwork creation. The other metrics evaluated, mean absolute percent error and the structural similarity index, may have a great potential for enhancing other types of transportation system analysis. However, these metrics had a limited potential to identify the important aspects of subarea input variation. The root means square error, or RMSE, of the induced subnetwork origin-destination matrices was determined to be the most effective measure of rerouting occurring at the boundary of the subnetwork. Now the investigation focused on developing a means of accurately comparing this metric to determine the differences in impact scenario and base scenario inputs to the subnetwork.

For typical traffic simulation, use of multiple simulations (replicate runs) of a network to account for real world variability, which is also imperative to the subarea inputs. Incorporating the variation prevents the use of a general quantification of the input deviations. Instead, a more appropriate statistical technique like a two sample equal means test can account for the randomness in simulation results. Using the average of the induced subarea demand in the base scenario as the true value for the input, each impact scenario and base scenario simulation result was compared to the true values by calculating the RMSE. These RMSE values are then compared using a two sample equal means test for the base scenario sample and the impact scenario sample. This test accomplishes the need for determining if the rerouting in the impact scenario simulation is statistically significantly greater than the base scenario simulations. Applying the equal means test across different subnetwork sizes reveals that increasing the subnetwork size increases the capability of the subnetwork to capture the rerouting and congestion within its extents. At this subnetwork size, when the majority of rerouting is captured, the sufficient subnetwork size has been identified. This method captured the trends that were expected with an increase in subnetwork size. This verification was an essential step, but a new method that could equally predict a sufficient subnetwork without the need for multiple simulations was desired.

### 7.3 Summary of Prediction Technique

In order to reduce the burden of reproducing the subnetwork analysis performed for the comparison technique, it was hypothesized that a statistical prediction model could enhance the comparison analysis. This would also enable dynamic traffic assignment users to diagnose the inputs that have an impact on their analysis. The variables used to make recommendations for the capacity analysis, the capacity reduction and number of links impacted, were included in an initial model to verify the applicability of this direction for the analysis. Subnetwork size was included in the model as the primary variable of study, so that a user may input a variety of subnetwork sizes in order to predict the expected RMSE of the boundary demand. Promising results from a preliminary model were used as justification for examining further variables that have a significant influence on DTA rerouting.

Testing a variety of model specifications revealed the important variables for investigating an impact analysis. The key additional data that was provided for the model was based on link information from the full network base scenario. This included the original capacity and volume to capacity ratio of the impacted links for the scenario. These values incorporate into the rerouting prediction model the amount of traffic that the impact location can handle and the level of congestion in the area prior to network alterations. Another important factor that was included was the interaction term between subnetwork size and capacity reduction. This indicates that for some small levels of capacity reduction, there may be limited change in rerouting as subnetwork size increases. This is intuitive because the variation in rerouting would be expected to be constant in the near base condition (small impact) scenario. After generating an effective model for predicting rerouting in the impact scenario induced subarea demand, a method was sought for implementing the model to generate a sufficient subnetwork.

The comparison method was capable of determining when the base scenario subnetwork demand was statistically similar to the impact scenario subnetwork demand. Therefore, if two regression models could accurately predict when the base and impact scenario are no longer very different then this could provide similar results to the comparison model. When the final impact model was selected, a base model was generated using the same variables as the impact model, excluding those that applied only to the impact. These final base and impact models were found to accurately predict the amount of rerouting that occurred at the subarea boundaries in each scenario. More importantly, the intersection point of the lines generated by the two models predicted similar recommendations as the comparison technique. This now provides a means for automating and generalizing the sufficient subnetwork size recommendations.

### 7.4 CONCLUDING REMARKS AND RECOMMENDATIONS

This dissertation developed a methodology based on case study data. This is an appropriate approach for this type of analysis for a few reasons. Primarily, case study theory is most suited to new research areas (Eisenhardt, 1989). Since subnetwork analysis has been primarily ad hoc until this contribution, the testability of the empirical validity is essential for building a foundation for subarea methodology. This study began with a "within case" study, Austin, to generate the theory and used a "cross case" study, Dallas, to address the validity. The theoretical sampling, ten simulations with a random seed of representative scenarios (the golden rule for traffic simulations), allowed for the examination of useful, practical scenarios. Two simulations are also commonly used in practice to reduce the burden of time required to simulate traffic ten times. Also, the cross case study fixed the seed for random number generation to addressing the exact difference between base and impact scenarios. Application of the case study approach has identified a verified, transferable method of analysis.

If a user is concerned about the validity of this modeling approach for a particular region or scenario, then the framework for building the model can be reapplied. This technique is akin to the Mechanistic-Empirical Pavement Design Guide, where models are generated for predicting results and default coefficients are provided for the parameters. However, it is acknowledged that differences occur between different regions and suggests that calibrating the model to the parameters of the region can improve analysis. The coefficients presented here are a good default for practitioners, but if desired the methodology may be performed to account for potential differences.

This method is both a common technique in transportation engineering, but also addresses a common effect seen in transportation networks. In geographic terms, things that are closer together tend to have a greater impact than things that are farther away. This magnitude of effect decreasing with distance has been demonstrated in transportation when looking at environmental effects of a rail corridor (Wang, 2014). It is suspected that in a perfect grid network impacts may take the form of a three dimensional normal distribution decreasing as the distance from impact increases. These effects may become less symmetrical as the network becomes less homogeneous. However, the regression model proposed in this dissertation attempts to capture this trend with distance (connected order size).

There are many applications that can extend from this analysis. Particularly, a further investigation of the network topology metrics may improve the understanding of subnetwork and multi-resolution analysis. It is important to characterize the network when presenting measures of effectiveness because an equal level of service may actually be dramatically different in terms of desired performance. An extremely sparse network cannot expect to have great levels of service during major roadway construction since there exist few alternative routes. Network uniqueness is an obstacle due to the large number of factors, but creating some method of classification can help. Results in this dissertation also indicate that network size may have some influence on convergence. Perhaps, convergence is more of a model resolution issue than it is a matter of algorithms reducing the number of iterations needed to provide accurate results. It should be addressed whether examining subareas of a regional network may be used to parallelize dynamic traffic assignment more effectively and generate results with a greater coverage more quickly.

# SECTION 3 SUMMARY: DESTINATION

This dissertation has arrived at its final destination by effectively applying tools and techniques to enhance the understanding of subnetwork analysis. It accomplished this through a new application of common transportation analysis tools to measure the spatial and temporal extents of congestion effects in traffic simulation. This contribution is substantial to the development of multi-resolution analysis because it offers insight into the important causes and effects to consider when analyzing the geographic scope of analysis. The tools and techniques developed in this dissertation may be used by practitioners to understand and defend their decisions when generating a network for traffic simulation. In transportation planning the destination zone is the ultimate goal for each vehicle user, and providing a means for implementation is the ultimate goal of this dissertation. Recommendations for use of this method and what can be done in the future are also included in this destination section.

The contribution of this dissertation has many applications for future research and development of dynamic traffic assignment and transportation simulation. Primarily, major strides have been made to improve DTA algorithms such as time-dependent shortest path calculations and route choice algorithms. Using a real subnetwork as a test network for these algorithms will enhance their development. Other algorithms that may cause a dramatic increase in simulation time could become more tractable from the use of subnetworks. Managing simulation time will become more important as more sophisticated levels of modeling are introduced including: variability of link performance, multimodal transportation options, intersection treatment, and new or modified traffic flow models. This dissertation has documented the considerations one must make when implementing a subnetwork procedure on network modeling.

## References

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., & Mishalani, R. (1998). DynaMIT: a simulation-based system for traffic prediction. In DACCORS Short Term Forecasting Workshop, The Netherlands.

Binkowski, S. E., & Hicks, J. E. (2013). Development and Use of a 700-square mile DTA Model for Corridor Maintenance of Traffic Decision-Making. Transportation Research Board 93rd Annual Meeting.

Bierlaire, M., Crittin, F., (2004) An Efficient Algorithm for Real-Time Estimation and Prediction of Dynamic OD Tables. Operations Research 52(1):116-127. http://dx.doi.org/10.1287/opre.1030.0071

Boxill, S. A. (2005). The Integration of GIS and Transportation Modeling: A State-of-the-Practice Review (No. SWUTC/05/473700-00043-1).

Boyles, S. D., & Waller, S. T. (2011). Optimal information location for adaptive routing. Networks and Spatial Economics, 11(2), 233-254.

Bringardner, J. W. (2012). Developing a GIS-based intersection traffic control planning tool.

Bringardner, J. W., Gemar, M. D., Boyles, S. D., & Machemehl, R. B. (2014). Establishing the Variation of Dynamic Traffic Assignment Results Using Subnetwork Origin-Destination Matrices. In Transportation Research Board 93rd Annual Meeting (No. 14-3107).

Burghout, W., Koutsopoulos, H. N., & Andreasson, I. (2005). Hybrid mesoscopic-microscopic traffic simulation. Transportation Research Record: Journal of the Transportation Research Board, 1934(1), 218-255.

Carsjens, G. J., & Ligtenberg, A. (2007). A GIS-based support tool for sustainable spatial planning in metropolitan areas. Landscape and urban planning, 80(1), 72-83.

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(3), 241-256.

Chen, B.Y., Lam, W.H., Sumalee, A., Li, Q., Li, Z.C. (2012). Vulnerability analysis for large-scale and congested road networks with demand uncertainty. Transportation Research Part A 46 (3), 501-516.

Chien, I. J. S., Mouskos, K., & Ziliaskopoulos, A. K. (2005). Development of a Simulation/assignment Model for the NJDOT I-80 ITS Priority Corridor. Department of Transportation, the State of New Jersey.

Chiu, Y. C., & Zheng, H. (2007). Real-time mobilization decisions for multi-priority emergency response resources and evacuation groups: model formulation and solution. Transportation Research Part E: Logistics and Transportation Review, 43(6), 710-736.

Chiu, Y. C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, T., Hicks, J. (2010). DTA Primer, Network Modeling Committee (ADB30) of the Transportation Research Board, Washington, DC.

Chiu, Y.C., November 21 (2013). Yi-Chang Chiu: Solving traffic congestion through cooperation and reward [Video File]. Retrieved from http://tedxtalks.ted.com/video/Solving-traffic-congestion-thro

Choi, K., Jayakrishnan, R., Kim, H., Yang, I., & Lee, J. (2009). Dynamic Origin-Destination Estimation Using Dynamic Traffic Simulation Model in an Urban Arterial Corridor. Transportation Research Record: Journal of the Transportation Research Board, 2133(1), 133-141.

Cools, M., Moons, E., Wets, G. (2010). Assessing the Quality of Origin-Destination Matrices Derived from Activity Travel Surveys. Transportation Research Record: Journal of the Transportation Research Board 2183 (1), 49-59.

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

Daganzo, C. F., (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B 28 (4), 269-287.

Daganzo, C. F., (1995). The cell transmission model, part II: network traffic. Transportation Research Part B 29 (2), 79-93.

Deng, Q., & Cheng, L. (2013). Research Review of Origin-Destination Trip Demand Estimation for Subnetwork Analysis. Procedia-Social and Behavioral Sciences, 96, 1485-1493.

Djukic, T., Hoogendoorn, S., Van Lint, H. (2013). Reliability Assessment of Dynamic OD Estimation Methods Based on Structural Similarity Index. Transportation Research Board 92nd Annual Meeting. (No. 13-4851).

Evangelidis, K., Papaioannou, P., & Panayotakopoulos, D. (2005). Considering Network Demand Issues in GIS Transportation Data Modeling. In Proceedings in electronic format (CD-ROM) of the 2005 ESRI International User Conference.

FHWA GIS-T http://www.gis.fhwa.dot.gov/

Florian, M., Mahut, M., & Tremblay, N. (2008). Application of a simulation-based dynamic traffic assignment model. European Journal of Operational Research, 189(3), 1381-1392.

Galaand, C., & Scotton, P. (1996). U.S. Patent No. 5,495,479. Method and apparatus for an automatic decomposition of a network topology into a backbone and subareas. Washington, DC: U.S. Patent and Trademark Office.

Gartner, N. H., Golden, B. L., & Wong, R. T. (1976). Modeling and Optimization for Transportation Systems Planning and Operations.

Gartner, N. H., Messer C.J., Rathi, A.K. (2005). Revised monograph on traffic flow theory: a state-of-the-art report. Special Report by the Transportation Research Board of the National Research Council. http://www.fhwa.dot.gov/publications/research/operations/tft/

Gemar, M. D. (2013). Subnetwork analysis for dynamic traffic assignment: methodology and application.

Gemar, M., Bringardner, J., Boyles, S., Machemehl, R. (2014). Subnetwork Analysis for Dynamic Traffic Assignment Models: A Strategy for Estimating Demand at Subnetwork Boundaries. Transportation Research Record: Journal of the Transportation Research Board, in press.

Hadi, M., Shabanian, S., Ozen, H., Xiao, Y., Doherty, M., Segovia, C., & Ham, H. (2013). Application of Dynamic Traffic Assignment to Advanced Managed Lane Modeling.

Han, K., Piccoli, B., Friesz, T. L., Yao, T. (2012). A continuous-time link-based kinematic wave model for dynamic traffic networks. arXiv preprint arXiv:1208.5141.

Hardy, M., Wunderlich, K. (2009). Traffic Analysis Tools Volume IX: Work Zone Modeling and Simulation—A Guide for Analysts, US Department of Transportation: FHWA-HOP-09-001. http://ops.fhwa.dot.gov/trafficanalysistools/index.htm

Hicks, J. E. (2008). Dynamic Traffic Assignment Model Breakdown. In Conference Proceedings 42, Innovations in Travel Demand Modeling, Summary of a Conference (Vol. 2, pp. 101-108).

Hossack, S., & Ortega, J. (2012). Leveraging Web-Based GIS for Managing Transportation Network Information. In Transportation Research Board 91st Annual Meeting (No. 12-0884).

Jeannotte, K., Chandra, A., Alexiadis, V., Skabardonis, A. (2004). Traffic Analysis Toolbox Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools (No. FHWA-HRT-04-039).

Jenelius, E., Petersen, T., & Mattsson, L. G. (2006). Importance and exposure in road network vulnerability analysis. Transportation Research Part A: Policy and Practice, 40(7), 537-560.

Jenelius, E. (2009). Network structure and travel patterns: explaining the geographical disparities of road network vulnerability. Journal of Transport Geography, 17(3), 234-244.

Jin, P. J., Boyles, S. D. (2014). The Travel Time Transmission Model for Dynamic Network Loading. In Transportation Research Board 93rd Annual Meeting (No. 14-5718).

Kamga, C. N., Mouskos, K. C., Paaswell, R. E., (2011). A methodology to estimate travel time using dynamic traffic assignment (DTA) under incident conditions. Transportation Research Part C 19 (6), 1215-1224.

Khalesian, M., Pahlavani, P., & Delavar, M. R. (2009). A GIS-based Traffic Control Strategy Planning at Urban Intersections. IJCSNS, 9(1), 166.

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(5), 843-854.

Larsson, T., Lundgren, J. T., Rydergren, C., & Patriksson, M. (2001). Most likely traffic equilibrium route flows analysis and computation. In Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Models (pp. 129-159). Springer US.

Li, C., Ying, X., Honglei, Z., & Xiang-lei, Y. (2010). Dynamic division about traffic control sub-area based on back propagation neural network. In Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2010 2nd International Conference on (Vol. 2, pp. 22-25). IEEE.

Lillesand, T.M., Kiefer, R.W., Chipman, J.W. (2008). Remote sensing and image interpretation. (No. Ed. 6). John Wiley & Sons Ltd.

Lowry, J. H., & Lowry, M. B. (2014). Comparing spatial metrics that quantify urban form. Computers, Environment and Urban Systems, 44, 59-67.

Mahmassani, H. S., Hu, T., & Jayakrishnan, R. (1995). Dynamic traffic assignment and simulation for advanced network informatics (DYNASMART). Urban traffic networks: Dynamic flow modeling and control. Springer, Berlin/New York.

Mahmassani, H. S. (2001). Dynamic network traffic assignment and simulation methodology for advanced system management applications. Networks and Spatial Economics, 1 (3-4), 267-292.

Mahut, M., Florian, M., Tremblay, N., Campbell, M., Patman, D., & McDaniel, Z. K. (2004). Calibration and application of a simulation-based dynamic traffic assignment model. Transportation Research Record: Journal of the Transportation Research Board, 1876(1), 101-111.

Martin, W. A., & McGuckin, N. A. (1998). Travel estimation techniques for urban planning (No. 365). Washington, DC: National Academy Press.

Marzano, V., Papola, A., Simonelli, F. (2008). Effectiveness of Origin-Destination Matrix Correction Procedure Using Traffic Counts. Transportation Research Record: Journal of the Transportation Research Board, 2085 (1), 57-66.

Melson, C. L. (2013). Improvements and extensions of dynamic traffic assignment in transportation planning.

Mitsakis, E., Salanova, J. M., Giannopoulos, G. (2011). Combined dynamic traffic assignment and urban traffic control models. Procedia-Social and Behavioral Sciences, 20, 427-436.

National States Geographic Information Council (2011), "Transportation for the Nation Strategic Plan," The US Department of Transportation Geospatial Information Officer.

Nezamuddin and S. D. Boyles (2014). A continuous DUE algorithm using the link transmission model. In press, Networks and Spatial Economics.

Ni, D. (2011). Multiscale modeling of traffic flow. Mathematica Aeterna, 1(1), 27-54.

O'Sullivan, D., & Unwin, D. J. (2010). Geographic information analysis.

Paz, A., Chiu, Y. C. (2011). Adaptive Traffic Control for Large-Scale Dynamic Traffic Assignment Applications. Transportation Research Record: Journal of the Transportation Research Board, 2263 (1), 103-112.

Peeta, S., Ziliaskopoulos, A.K. (2001). Foundations of dynamic traffic assignment: the past, the present and the future. Networks and Spatial Economics 1 (3-4), 233-265.

Pesti, G., Chu, C. L., Balke, K., Zeng, X., Shelton, J., & Chaudhary, N. (2010). Regional Impact Of Roadway Construction On Traffic And Border Crossing Operations In The El Paso Region. Center for International Intelligent Transportation Research, Texas Transportation Institute.

Pool, C. M. (2013). Enhancing the practical usability of dynamic traffic assignment.

Qian, Z., Zhang, H. M. (2012). On centroid connectors in static traffic assignment: Their effects on flow patterns and how to optimize their selections. Transportation Research Part B: Methodological, 46(10), 1489-1503.

Quiroga, C., & Koncz, N. (2008). Engineering Design Data Management-Practices and Framework Development (No. FHWA/TX-08/0-5246-1).

Quiroga, C., Li, Y., Koncz, N., & Overman, J. (2009). Analysis and Integration of Spatial Data for Transportation Planning (No. FHWA/TX-09/0-5696-1). Texas Transportation Institute, Texas A&M University System.

Schroeder, B., Rouphail, N., Williams, B., Hajbabaie, A., Jia, A., Sajjadi, S., Aghdashi, B., and Narron, B. (2014). NCDOT Project 2012-36. Work Zone Traffic Analysis & Impact Assessment.

Scopatz, B., Lefler, N., & Eccles, K. (2013). Assessment of the Geographic Information Systems'(GIS) Needs and Obstacles in Traffic Safety (No. FHWA-HRT-13-096).

Sheffi, Y. (1984). Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Englewood Cliffs, NJ: Prentice-Hall.

Sisiopiku, V. P., Li, X., Mouskos, K. C., Kamga, C., Barrett, C., & Abro, A. M. (2007). Dynamic traffic assignment modeling for incident management. Transportation Research record: Journal of the transportation research board, 1994(1), 110-116.

Sisiopiku, V. P., Sullivan, A. J., Abro, A. M., Shinouda, M., Mouskos, K., & Barrett, C. (2010). Development of a Dynamic Traffic Assignment Model to Evaluate Lane-Reversal Plans for I-65 (No. UTCA Report Number 07408). University Transportation Center for Alabama.

Sloboden, J., Lewis, J., Alexiadis, V., Chiu, Y. C., & Nava, E. (2012). Traffic Analysis Toolbox Volume XIV: Guidebook on the Utilization of Dynamic Traffic Assignment in Modeling (No. FHWA-HOP-13-015).

Sundaram, S., Koutsopoulos, H. N., Ben-Akiva, M., Antoniou, C., & Balakrishna, R. (2011). Simulation-based dynamic traffic assignment for short-term planning applications. Simulation Modelling Practice and Theory, 19(1), 450-462.

Sullivan, J. L., Novak, D. C., Aultman-Hall, L., Scott, D. M., (2010). Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach. Transportation Research Part A 44 (5), 323-336.

Tamminga, G., van den Brink, L., Van Lint, H., Stoter, J., & Hoogendoorn, S. (2013). Toward GIS-Compliant Data Structures for Traffic and Transportation Models. In Transportation Research Board 92nd Annual Meeting (No. 13-2455).

Taylor, M. A., Sekhar, S. V., & D'Este, G. M. (2006). Application of accessibility based methods for vulnerability analysis of strategic road networks. Networks and Spatial Economics, 6(3-4), 267-291.

Taylor, M. A. (2012). Remoteness and accessibility in the vulnerability analysis of regional road networks. Transportation research part A: policy and practice, 46(5), 761-771.

Tsekeris, T., & Geroliminis, N. (2013). City size, network structure and traffic congestion. Journal of Urban Economics, 76, 1-14.

Traffic Analysis Tools Program, Federal Highway Administration, Washington, DC, (2014). Accessed online: March 29, 2014. http://ops.fhwa.dot.gov/trafficanalysistools/index.htm.

Waller, S. T., Duthie, J. C., Nezamuddin, N., Juri, N. R., Rambha, T., Melson, C., & Kumar, R. (2012). Investigating Regional Dynamic Traffic Assignment Modeling for Improved Bottleneck Analysis (No. 0-6657).

Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. Image Processing, IEEE Transactions on 13 (4), 600-612.

Wang, X. (2005). Integrating GIS, simulation models, and visualization in traffic impact analysis. Computers, Environment and Urban Systems, 29(4), 471-496.

Xie, C., Kockelman, K. M., & Waller, S. T. (2010). Maximum entropy method for subnetwork origin-destination trip matrix estimation. Transportation Research Record: Journal of the Transportation Research Board, 2196(1), 111-119.

Yazici, M. A., Kamga, C., & Mouskos, K. C. (2012). Analysis of Travel Time Reliability in New York City Based on Day-of-Week and Time-of-Day Periods. Transportation Research Record: Journal of the Transportation Research Board, 2308(1), 83-95.

Yperman, I. (2007). The Link Transmission Model for Dynamic Network Loading. PhD Dissertation, Katholieke Universiteit Leuven, Leuven, Belgium.

Zhang, M., & Kukadia, N. (2005). Metrics of urban form and the modifiable areal unit problem. Transportation Research Record: Journal of the Transportation Research Board, 1902(1), 71-79.

Zhou, X., Qin, X., & Mahmassani, H. S. (2003). Dynamic origin-destination demand estimation with multiday link traffic counts for planning applications. Transportation Research Record: Journal of the Transportation Research Board, 1831(1), 30-38.

Zhou, X., Erdogan, S., Mahmassani, H. S. (2006). Dynamic Origin-Destination Trip Demand Estimation for Subarea Analysis. Transportation Research Record: Journal of the Transportation Research Board, 1964 (1), 176-184.

Zhou, Y., Liang, H., Wu, J., & Gao, L. (2009). Research on integration of microscopic traffic simulation and GIS. In Service Operations, Logistics and Informatics, 2009. SOLI'09. IEEE/INFORMS International Conference on (pp. 657-662). IEEE.

Ziliaskopoulos, A. K., & Waller, S. T., (2000). An Internet-based geographic information system that integrates data, models and users for transportation applications. Transportation Research Part C: Emerging Technologies, 8(1), 427-444.

Ziliaskopoulos, A. K., Waller, S. T., Li, Y., & Byram, M. (2004). Large-scale dynamic traffic assignment: Implementation issues and computational analysis. Journal of Transportation Engineering, 130(5), 585-593.