The Thesis Committee for Rachel Michelle James
Certifies that this is the approved version of the following thesis:

Data-Driven Placement of Centroid Connectors in Dynamic Traffic Assignment

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor: ______________________________
Stephen D. Boyles

____
Mason D. Gemar

Data-Driven Placement of Centroid Connectors in Dynamic Traffic Assignment

by

Rachel Michelle James, B.S.C.E.

Thesis

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

Master of Science in Engineering

The University of Texas at Austin

August 2016

Dedication

This thesis is dedicated to Joe and Robin Hale. My earliest vivid childhood memory is of my grandparents instructing me that education was my one-way ticket to a better life, which is advice I have carried near my heart throughout my academic studies. My Popaw and Momaw are the greatest blessing I could have ever asked for. By instilling values in me—like a love of learning, compassion for others, and grit, determination, and maturity far beyond my years—they gave this little girl from rural West Virginia wings to fly far beyond what should have been possible for me to ever achieve.

My Popaw was with me through the beginning of this chapter of my life, but is heartbreakingly not here to see the conclusion. Thank you for all you've done for me, Popaw. I love you and am so thankful to be your granddaughter. I'm so sorry these wings didn't get me home to see you one last time.

Acknowledgements

I would like to start off my acknowledgements with a nod to the faculty and staff at the Statler College of Engineering at West Virginia University, specifically Dr. Avinash Unnikrishnan, Dr. David Martinelli, and Dr. John Zaniewski, for helping me to grow into the professional I am today and for encouraging me to pursue graduate studies. I'm incredibly thankful for the opportunities provided by Dr. Avi during the three years I spent in his lab as an undergrad research assistant—his mentorship means the world to me.

Next, I want to thank the faculty I have encountered in my short time at UT Austin—the opportunity to learn from you has been second to none. In particular, I would like to thank my research advisor, Dr. Stephen Boyles, for taking a chance on me and providing me the opportunity to research and grow in his lab at UT Austin. Thank you, Dr. Boyles, for your advice and support on this particular research effort. I would also like to thank Dr. Nati Ruiz Juri, Dr. Mason Gemar, and Dr. Jen Duthie for my time spent at the Networks Modeling Center, as well as their support and mentorship with this project. I am so incredibly grateful to my labmate, Ehsan, for extending his previous code for centroid connector placement to allow for the inclusion of parcel data. Without you, Ehsan, this research effort would have not been nearly as efficient and I appreciate your time and effort. I would also like to thank the remainder of my research group, for providing base code, helping to brainstorm ideas, or providing moral support. The last thesis related shout out belongs to Kasun for giving me insanely quick feedback on this document and for always being willing to give me advice.

One of my favorite things about graduate school has been the opportunity to get to know so many incredible people. In particular, I want to thank Dr. Meredith Celeblak, Venktesh, Kristie, John, Priyadarshan, and Felipe for all of the support, fun, and friendship had over the last two years. I'd also like to thank my Austin "FitFam"—notably Amanda, Dr. Steve, and Sara—for helping me to fall in love with fitness. My lovely roommate, Michelle, also deserves a shout out for being such an awesome person to live with, as well as for putting up with me, Minnie, and Lily. Lastly, I want to thank the Vestri family—Greg, Jean, Rachel, Ariel, Drew, Addie, and Jacob—for essentially unofficially adopting me and being such a large part of my Austin experience. I'm so thankful grad school gave us an opportunity to reconnect and am so thankful you're such a large part of my life.

Last but not least, I would like to thank my grandparents, Joe and Robin Hale, and my siblings, Adam, Brian, Michael, Isaac, and Elijah, for being a constant source of love, support, and inspiration. You are my motivation and the reason I have made it this far. I'd also like to thank my sweet Ian, for being an infinite source of love and support, and his wonderful family, Roger and Mary Ann Murray and their wonderful kids and grandkids, for taking me in as if I were one of their own.

I am so blessed to have each and every one of you in my life. Thank you for making this possible.

Abstract

Data-Driven Placement of Centroid Connectors in Dynamic Traffic Assignment

Rachel Michelle James, M.S.E

The University of Texas at Austin, 2016

Supervisor: Stephen D. Boyles

Recent technological advances allows transportation engineering professions to collect, share, and handle unprecedented quantities of data, which has the potential to transform current transportation planning paradigms. In the immediate future, data can be used to improve the precision and capabilities of existing transportation network modeling frameworks. Parcel data is a large, readily available data source that represents the location of public lands, businesses, and residences and is frequently used by government and businesses for land use and zoning decisions. This thesis looks at the viability of using parcel data to inform static traffic assignment (STA) and dynamic traffic assignment (DTA) connector placement in a medium sized network in the Austin, TX region.

Simulation-based DTA models are particularly sensitive to the topological detail of the traffic network, including the location of centroid connectors. Traditional centroid connector placement strategies may lead to excessive congestion and unrealistic traffic patterns, while manual network refinement is prohibitive in large regional models. In this

thesis, parcel-level data is used to both allocate travel demand between two sub-regions in each considered traffic analysis zone and to select appropriate nodes for the centroid connector placement. Numerical experiments suggest that the proposed approach better approximates both corridor travel times and traffic counts throughout the network, with improvements of more than 40 percent in travel time estimation accuracy, and 12 percent in traffic count estimation. Additionally, the scenarios that best matched count and travel time data were the scenarios that had the highest average parcel density per entry/exit node, indicating that parcel data is an acceptable proxy for high demand points in the network.

When applied in STA, the results were not quite as promising. Although this methodology was able to improve the utilization of lower capacity links, the results ultimately did not better resemble volume count data. However, this does represent a simple, transparent, and data-driven approach for centroid connector placement in static traffic assignment that performs as well as traditional methods.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background

One critical component of transportation planning is travel demand modeling, which, simply put, is a mathematical model of the supply and demand for travel in an urban environment. Though the state-of-the-art practice in travel demand modeling is an activity based approach, the most common modeling paradigm in practice is the “four step” model (McNally, 2007). The four step model is an aggregated trip-based methodology that yields the number of trips between each origin-destination pair within a study area. Transportation professionals define origins and destinations within planning models through a unit of geography known as a traffic analysis zone (TAZ). The urban area is partitioned into a number of TAZs that can range in size from a city block to an entire neighborhood depending on the scale of the model. Each TAZ is represented by a centroid, which is an aggregation of all of the real origins and destinations within a TAZ and typically physically located at the geometric center of gravity of the TAZ (Sheffi, 1985). The centroid is connected to the model network via a number of artificial links, called centroid connectors, which represents all of the lower capacity roads within the TAZ that are not modeled explicitly on the network.

Once the TAZ structure and the network representation have been established, the four step trip-based modeling approach involves trip generation, trip distribution, modal split, and traffic assignment modules (see Figure 1.1). In trip generation, the aggregate number of trips produced at and attracted to a TAZ is estimated as a function of land use, socio-economic variables, and trip purpose. In trip distribution, origin-destination pairs are created by determining the number of inter-zonal trips as a function of impedance, typically through a gravity model. The third step, modal split, determines the proportion of trips made by various modes, often determined by a logit model. The last step of the four step model is traffic assignment, where the expected route of each trip/vehicle is found via the principle of user equilibrium; this step is the primary research focus of this thesis.

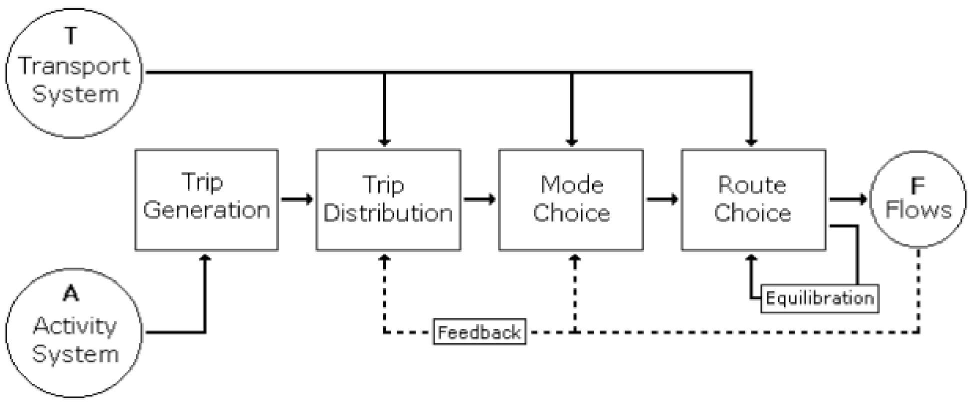


Figure 1.1: The Four Step Travel Demand Model (McNally, 2007)

During this last step, all of the trips found to travel between zones by the car mode during the prior three steps are assigned to the network to observe resulting travel (link and path flows) patterns. The output, the predicted route choice of vehicles, is used to develop performance metrics that assist in the evaluation of alternatives and to inform transportation planning and policies. There are two primary categories of traffic assignment models: static and dynamic. Though dynamic traffic assignment (DTA) is the state-of-the-art for traffic assignment—as it more realistically captures congestion propagation, queue spillback, and network delay—the mathematical properties of static traffic assignment (STA) make it useful for researchers and in practical planning scenarios where data inputs are not known with certainty. Thus, though some view dynamic traffic assignment as a superior tool to static traffic assignment, they are different tools for different problems—depending on model purpose, desired outcomes, and availability of expertise and funding. Regardless of the nature of the modeling approach, the organization of TAZs, definition of centroids, and the placement of centroid connectors provide the foundation for any model and are pivotal to the success of any venture.

This thesis focuses on the placement of centroid connectors and their impact on the results of traffic assignment. It has been observed that one unintended consequence of the network abstraction and the user equilibrium behavior assumption in traffic assignment is that simulated vehicles can utilize centroid connectors in a manner that is behaviorally inconsistent with reality (e.g. utilize the connector as a “free ride” to skip over lower capacity roadways within a TAZ, systematically overloading higher capacity/higher speed links). As stated by Sheffi, “The issue of network representation is as much an art as a science; practice and experience are required to carry it out successfully” (Sheffi, 1985, pp. 16). This thesis seeks to make the centroid connector placement task in network representation determination more systematic by leveraging readily available parcel data to streamline the calibration of static and dynamic traffic assignment models.

1.2 MOTIVATION

The use of advanced transportation planning models allows decision makers to better understand the spatial and temporal aspects of transportation systems’ performance. Among these, dynamic traffic assignment models, have been increasingly used in practice due to their ability to capture the propagation of congestion and the impact of a variety of traffic control and management strategies (Chiu, Bottom, Mahut, Pax, Balakrishna, Waller, & Hick, 2011; Sloboden, Lewis, Alexiadis, Chiu, & Nava, 2012).

Advanced models often require more detailed inputs, including a more accurate representation of the transportation network. However, previous research suggests that improving the network representation without considering how travel demand is loaded onto the network may not yield the desired benefits (Jafari, Gemar, Ruiz Juri, & Duthie, 2015). Additionally, it's been shown numerically that the placement of centroid connectors plays a significant role in the results of both static and dynamic traffic assignment models (Jafari et al., 2015; Qian & Zhang, 2012). While traditional centroid placement techniques may not suffice towards this end, manual refinement of centroid connectors may be prohibitive in large regional models.

Additionally, the simulation nature of DTA places increased importance on the validation and calibration of a model. Calibration is defined as a resource intensive process by which the base year of the model is adjusted to ensure that the model output performance metrics are realistic and statistically representative of real-world data (Sloboden et al., 2012). Furthermore, DTA model validation and calibration requires access to a rich data set of traffic counts, travel times, and queue accumulation in order to be able to judge the model's ability to approximate real world conditions (Chiu et al., 2011).

According to *Dynamic Traffic Assignment: A Primer* (Chiu et al., 2011), calibration of DTA models is a major hurdle to large scale deployment of DTA because it is a time consuming process and requires extra attention to detail. Moreover, a recent survey of the 20 largest metropolitan planning organizations (MPO) in the United States, conducted by Cambridge Systematics, found that concerns about obtaining robust input data and issues surrounding the calibration of models are two of the largest barriers to widespread adoption of DTA (Cambridge Systematics, 2015).

New technology allows us to collect, share, and utilize unprecedented quantities of data. This is observed in practice primarily in the transportation operations sector, where real-time data is utilized in congestion pricing, variable speed limits, and active traffic management strategies. Though not as strongly supported in the literature, the availability of such data also has the potential to transform transportation planning paradigms. In the short term, data can be used to improve the precision and capabilities of existing modeling frameworks. This work exemplifies the use of readily-available built environment data to enhance practical implementations of DTA and STA models.

1.3 Contribution

This thesis makes two primary contributions to the research areas of static and dynamic traffic assignment. First, this thesis utilizes parcel-level data to inform an automated centroid connector placement methodology with the goal of producing more realistic network loading patterns based on the built environment surrounding existing network model nodes. Additionally, the location of parcel density in each TAZs is utilized to extend work by Jafari et al. (2015) found to improve the usage pattern of local streets in DTA models by dividing the TAZ demand between two artificial sub-regions comprising a TAZ; this effort utilizes parcel data to inform the division.

Though the detail oriented approach of connector placement via parcel data lends itself better to DTA mesoscopic analysis, this work presents a transparent methodology for centroid connector placement as a potential alternative strategy to traditional methods in static assignment. Perhaps, the larger contribution to STA literature is the use of real-world data to evaluate both the performance of the new methodology as well as the performance metrics found in the literature to compare and contrast the performance of centroid connector placement algorithms.

Numerical experiments, conducted on a medium-size network in the Austin, TX region, suggest that the proposed approach better approximates both corridor travel times and traffic counts for the dynamic traffic assignment model. Additionally, this application used in conjunction with static assignment was found to result in a significant increase in flow on lower capacity links and a marginal decrease in the flow on higher capacity functionally classified links. However, when field data were used as a validation metric, this methodology was not found to perform significantly better or worse than the original technique employed by the Capital Area Metropolitan Planning Organization (CAMPO) when creating the initial network structure. This suggests that looking for more reasonable flow patterns when analyzing the performance of static assignment centroid connector placement may not be a sufficient evaluation criterion.

1.4 Organization

The remainder of this thesis is organized as follows. Section 2 summarizes use of parcel data in planning and reviews literature discussing centroid connector placement in both static and dynamic traffic assignment. Section 3 briefly details algorithms and pseudocodes for dynamic and static traffic assignment before presenting the proposed methodology for the data-driven placement of centroid connectors. Section 4 provides detail on the experimental design and the scenarios modeled. Section 5 presents the numerical analyses, while Section 6 provides discussion of the results. Lastly, Section 7 offers concluding remarks and outlines possible future research directions.

Chapter 2: Literature Review

2.1 Introduction

The literature review begins with a brief refresher on static and dynamic traffic assignment and details VISTA (Visual Interactive System for Transport Algorithms), the DTA software utilized for this work. Next, the literature review briefly details the contrasting aspects of the traffic assignment methodologies relevant to this thesis—link performance functions versus link models based in fundamental traffic flow theory and their impact on the ability of the model to capture congestion propagation and the network fidelity and associated data requirements—and exhaustively details the literature on centroid connector placement. The literature review concludes with a synopsis of efforts to leverage big data in transportation, followed by a discussion on the current availability and utilization of parcel data.

2.2 Static Traffic Assignment

The goal of static traffic assignment (STA) is to find the flows on each link, x_{ij}, and the resulting travel time, $t_{ij}(x_{ij})$, at equilibrium conditions given specific network data (e.g. nodes, arcs, free flow speed, practical capacity, etc.). Static assignment uses link performance functions (LPF) to define a relationship between link flows and link travel times. One common link performance function is the Bureau of Public Roads (BPR) function (TRB, 1985) as shown below

$$t_{ij}(x_{ij}) = t_{ij}^{o}\left((1 + \alpha\left(\frac{x_{ij}}{c_{ij}}\right)^{\beta}\right) \tag{1}$$

where t_{ij}^{o} is the free flow travel time, x_{ij} is the flow on link ij, c_{ij} is the practical capacity of the link, and α and β are calibration factors (Equation 1).

In static traffic assignment, the fundamental behavioral assumption of route assignment is the principle of user equilibrium and follows Wardrop's first principle: for all origin-destination pairs, all used paths will have equal and minimal travel times (Wardrop, 1952). Static assignment can be formulated as an optimization model seeking to minimize the Beckmann function, as detailed in Chapter 3 (Beckmann, McGuire, & Winsten, 1956) and solved nearly exactly.

Static traffic assignment is extremely valuable to the transportation network analysis community because it is computational tractable and possesses robust mathematical properties. By Brouwer's theorem and variational inequalities, we know that user equilibrium exists. Furthermore, because the Beckmann function is strictly convex in link flows, the user equilibrium solution is unique in link flows.

Several variations of this problem are currently active areas of research. Some of these variations include elastic demand, destination choice, link interactions, perception error, and stochastic costs. In elastic demand and destination choice, the assumption that the origin-destination matrix is deterministic and a model input is relaxed and allowed to vary with congestion. To study link interactions, the assumption that the link travel time only depends on flows on its own link is relaxed. Lastly, for perception error and stochastic cost variations, the assumption that everyone is taking their known shortest path is relaxed.

Another active research topic in static assignment is the solution algorithm used to determine equilibrium conditions. Link-based algorithms, like the Method of Successive Averages (MSA) and Frank-Wolfe algorithm, were initially developed because they're economical with respect to computer memory; this is because they keep track of the link flows, a finite value, instead of the path flows. However, link-based algorithms are notoriously slow to converge, despite large steps toward equilibrium in the first few iterations of the algorithm. Conversely, path-based algorithms keep track of the path flows in each iteration. Though this greatly increases the memory required, it retains valuable information that can be utilized to converge to equilibrium more quickly than link-based algorithms. An example of a path-based algorithm is gradient projection. Bush-based algorithms, like Algorithm B, are the latest advances in assignment solution algorithms and offer the quick convergence to equilibrium, while requiring significantly less computational effort and memory than path-based algorithms.

Despite the quick convergence and robust mathematical properties, static assignment is unable to realistically model traffic flow and congestion. Some of the fundamental flaws with static assignment include its inability to model the time-dependent nature of traffic flow and the seemingly ‘arbitrary’ link performance functions used to model travel time as a function of flow. These two issues result in STA models that systematically underestimate the total system travel time, underestimate travel times on high capacity corridors, overestimate the number of travelers choosing high capacity routes as their shortest path, and underestimate the utilization of lower capacity links. These constraints motivated the study of DTA beginning in the 1970s (Chiu et al., 2011).

2.3 Dynamic Traffic Assignment

Static and dynamic traffic assignment have some very important parallels; however, where they differ, they do so intentionally. Though the field of DTA has not reached consensus on a single best practice methodology (Chiu et al., 2011), there are three primary elements that are required for a decision support tool to be classified as DTA. The first element is that there is some sort of model that accounts for how the fundamental variables of traffic flow (i.e. flow, density, speed, and vehicle number) change over time. This will ensure that congestion propagation through a network over time is explicitly modeled. Secondly, there must be some sort of concept of equilibrium with equilibrium route choice as an output. Finally, the equilibration must be based on experienced travel times, not instantaneous.

Dynamic traffic assignment was initially proposed as a tool for modeling traffic in the late 1970s by Merchant and Nemhauser (1978a, 1978b), when they proposed a non-linear non-convex mathematical program involving DTA with a single destination. DTA has matured significantly over the last 40 years. There are now two broad methodological categorizations of DTA approaches: analytical and simulation-based models. The three common analytical models seen through the literature are mathematical programming (Merchant & Nemhauser, 1978a; Carey, 1992; Ziliaskopoulos, 2000; Carey & Subrahmanian, 2000), optimal control (Friesz et al., 1989; Ran & Shimazaki, 1989a; Boyce et al., 1995; Ran et al., 1993), and variational inequality (Nagurney, 1998; Friesz et al., 1993; Chen & Hsueh, 1998) approaches. For more details on what each analytical approach entails, see Peeta and Ziliaskopoulos (2001). While analytical models have placed emphasis on maintaining the ability to derive theoretical insights, as with static traffic assignment, simulation-based models have focused on creating realistic models for practical deployment, regardless of the computational tractability.

Despite numerous research efforts to discover a closed form solution to dynamic user equilibrium (DUE), Carey (1992) proved the non-convex nature of constraints required in DTA and motivated a shift toward simulation-based solutions. The nature of this methodology places increased importance on the refinement of the network structure, to ensure it offers a real-world representation of the scenario, and the validation and calibration of the network, so that the results are consistent with travel patterns observed in reality (Sloboden et al., 2012). As mentioned above, with simulation-based DTA, the analytical properties of the problem formulation are lost in favor of the ability to more realistically model ill-behaved traffic flow in a practical manner. This is because a traffic simulator (be it microscopic, mesoscopic, or macroscopic) is used to model traffic flow and congestion propagation through the network (Mahmassani & Peeta, 1995; Jayakrishnan et al., 1994; Ben-Akiva et al., 1997).

2.4 Visual Interactive System for Transport Algorithms (VISTA)

This research effort uses VISTA as its DTA platform (Ziliaskopoulos & Waller, 2000). VISTA is a simulation-based approach to DTA. The VISTA framework iterates between two modules until convergence. In the "path generation" module, the time-dependent shortest path is found between each origin-destination pair at each departure time, a fixed percentage of vehicles are assigned to their identified shortest path, and the vehicles are simulated through the network, via the cell transmission model (Daganzo, 1994; Daganzo, 1995) to update travel costs. An iteration of "dynamic user equilibrium" determines the optimal percentage of vehicles to be shifted from their current path to the newly identified shortest path. These vehicles' trajectories are simulated through the network and the new path costs (and newest shortest paths) are identified. Convergence is evaluated after both modules are complete and the software terminates when travel times are found to be "sufficiently close" to equilibrium.

2.5 Comparison of Dynamic and Static Traffic Assignment

Static traffic assignment has a lot of redeeming qualities. The exact mathematical formulations are efficient and the provably correct solution methods of these models result in a fast, stable, and transparent methodology with unique equilibrium solutions. Additionally, STA is extremely valuable in practice when resources are limited, as it has much smaller input data requirements compared to DTA, or when input data are not known with a high degree of certainty, as STA is highly robust to input errors. However, this all comes at a high cost with respect to realism of results.

2.5.1 Propagation of Congestion and Route Choice Decisions

In addition to the obvious lack of temporal variation, one of the largest fundamental problems with static traffic assignment is with its use of link performance functions. By definition, capacity is the maximum flow rate that can be attained on a given segment of roadway (Figure 2.1); it occurs at the critical speed and critical density. It is a link's "tipping point", as any additional vehicle will cause congestion to set in and conditions will deteriorate with a reduction in speed and flow (Transportation Research Board, 2011). This definition and relationship is respected in DTA, as the network loading problem uses fundamental diagrams as inputs into the link models to determine the sending and receiving flows; realistically, the travel time found with link models increases slowly until the flow reaches capacity, at which point the travel time increases substantially, despite decreasing flow rates under congested conditions (Figure 2.2).

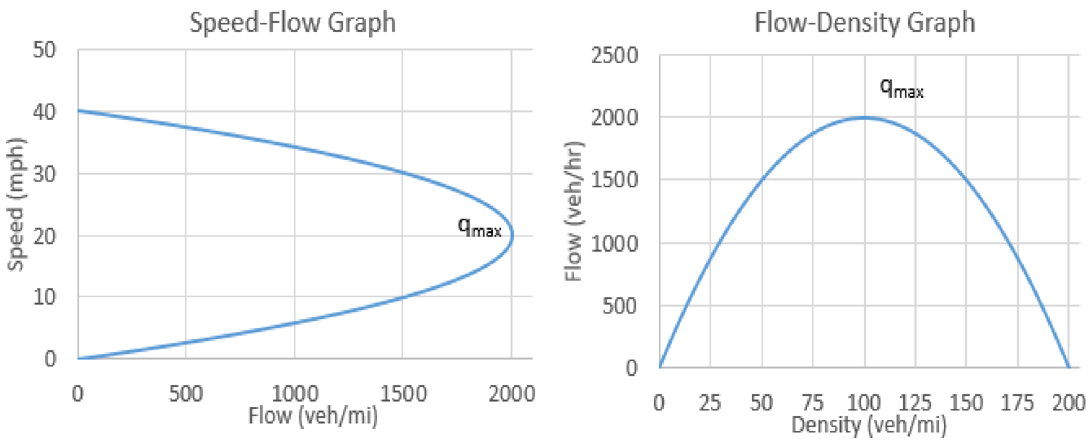


Figure 2.1: The relationship between speed, density and flow as defined by Greenshield

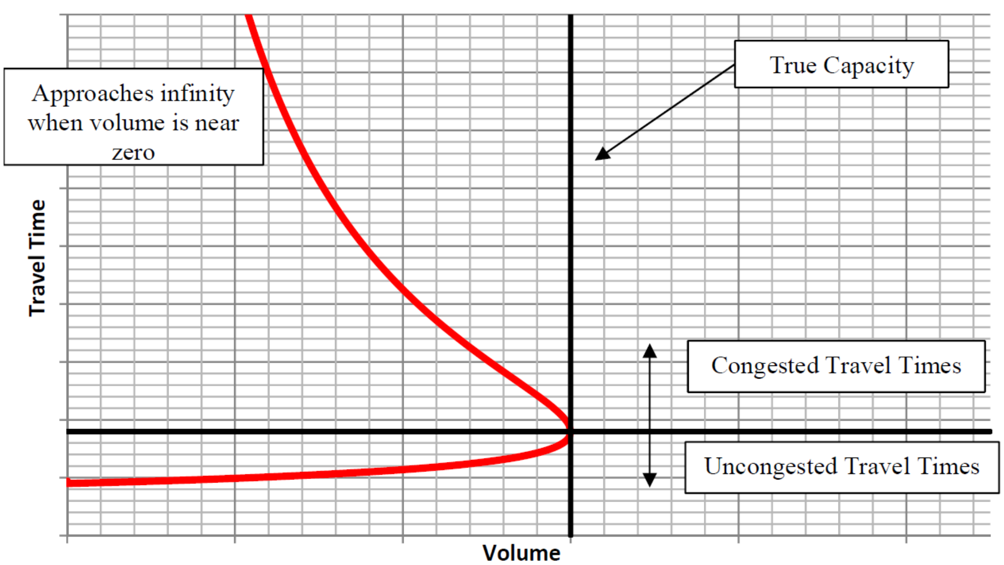


Figure 2.2: Travel times vs. flow in DTA (Duthie et al., 2013)

However, link performance functions do not enforce this relationship in STA. As shown in Figure 2.3, although travel times begin to increase once capacity is met, it does not increase in such a way that it influences people to find alternative routes when identifying their shortest path. In the literature, the relationship between travel time and flow via link performance functions has been found to cause a systematic overestimation of users on high-capacity route segments in static traffic assignment (Duthie et al., 2013; Boyles, Ukkusuri, Waller, & Kockelman, 2006). Two proposed solutions to this problem exist in the literature. One proposed solution to this problem involves using the "practical" capacity of a link, consistent with capacity at level-of-service (LOS) C or LOS D (Patriksson, 2015). Other researches have explored the importance of centroid connectors for achieving reasonable results in STA (see Section 2.6) (Friedrich & Galster, 2009; Qian & Zhang, 2012).

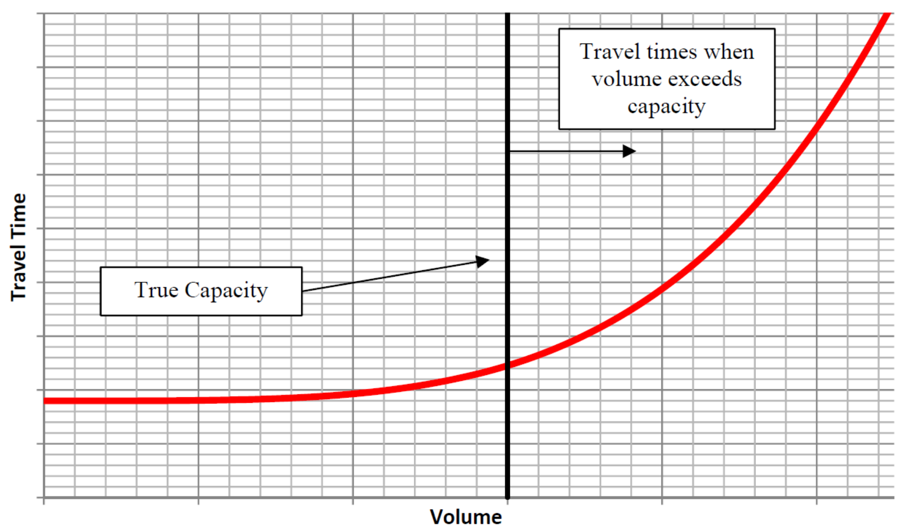


Figure 2.3: Travel times vs. flow in STA (Duthie et al., 2013)

2.5.2 Data Requirements and Necessary Level of Detail

As alluded to above, the simulation nature of DTA models is not as robust to input errors, placing increased significance on the importance of input data (Chiu et al., 2011). Thus, it is critical that the transportation network is portrayed accurately. In STA, required network data includes connectivity and data necessary for input into the link performance functions (e.g. free flow travel time, capacity, etc). In DTA, roadway geometrics become much more critical; information such as basic alignment and curvature, number of lanes, turning lane configuration, number of lanes, turning bay locations, intersection control devices, lengths of on- and off- ramps, operating speeds, etc. must be provided accurate to field conditions (Sloboden et al., 2012).

According to the DTA guidebook, the recommendations for level of detail and centroid connectors is as follows:

> "The model network needs to faithfully resemble the physical roadway network. It is critical that all the important intersection and roadway links of the study area are imported to the model. If links and intersections are omitted, they should generally belong to a roadway class that is at least one level down from the roadway class of the links on which the measures of effectiveness are collected. Local streets that provide access to adjacent properties and do not carry through traffic can be omitted and substituted with zone connectors." (Sloboden et al., 2012, p. 6-2).

Although there seems to be consensus in the literature about the proper level of abstraction in DTA models, not a lot of guidance is provided in the literature on where the abstracted local streets, or connectors, should be attached to the network (Jafari et al., 2015).

2.6 CENTROID CONNECTORS IN STATIC AND DYNAMIC TRAFFIC ASSIGNMENT

The consensus in the literature is that very little attention has been paid to the importance of centroid connector placement in network assignment models (Friedrich & Glaster, 2009; Qian & Zhang, 2012; Benezech & Leurent, 2013). In both *Dynamic*

Traffic Assignment: A Primer (Chiu et al., 2011) and the Federal Highway Administration *Guidebook on Dynamic Traffic Assignment Utilization* (Sloboden et al., 2012), the only guidance on centroid connector placement is that they should be placed at mid-block locations, not at signalized intersections, as the additional traffic movement may cause artificial congestion. In both Friedrich and Glaster (2009) and Qian and Zhang (2012), it was found that centroid connector placement is mostly proprietary information and left to the discretion of the software vendors. In the cases where information was available, they found that the connector nodes are typically the n closest nodes to the centroid (Benezech & Leurent, 2013), but they have also identified instances in the literature where connector nodes are assigned at random. Additionally, Qian and Zhang found recommendations that connector nodes should be selected along corridors with "intensive trip attractors/generators", but no guidance on how to identify those ideal nodes was provided (2012). Likewise, Friedrich and Glaster cite literature recommending that centroid connector nodes should be as close to natural access/egress nodes in the physical network, with no methodology for how that determination should be made for areas that have been deemed acceptable to be abstracted (2009).

Despite the gap in the literature, several authors have found that static and dynamic traffic assignment results are sensitive to how the networks are loaded. In the static world, Friedrich and Galster first brought this issue to light by stating that there is a discrepancy between the precision and accuracy of modeling the road network and the way demand is connected to the network in state-of-the-art planning models (2009). Through experimentation, they concluded that the lack of rules for how to establish connector nodes relies too much on the experience of the modeler, and thus negatively impacts the consistency and reliability of model outputs. They explored three methodologies for the selection of centroid connector access/egress points and concluded that there was no one optimal methodology (Friedrich & Galster, 2009).

Qian and Zhang expanded on this predecessor work and explored the behavior of results of static assignment on three network types: a synthetic grid network, a real world corridor, and the Sacramento area network (2012). Through sensitivity analysis, they found that traffic flow patterns have significant variations when centroid connector configurations are changed. Additionally, they found that adding additional connectors does not make the results more stable; instead, it often makes the results less realistic. However, too few connectors were found to result in artificial congestion on links where connectors are placed, thus indicating that more guidance on centroid connectors is necessary. These results indicate that static traffic assignment, which is often utilized for its robustness and stability of results, is unstable with respect to connector configuration and this often results in the underestimation of total corridor travel times (on higher capacity corridors) and average link flow. They proposed an optimization algorithm that seeks to minimize the maximum volume-to-capacity ratio of "characteristic links" (Qian and Zhang, 2012).

Given dynamic traffic assignment's increased sensitivity to network topology, it's not unreasonable to question if realistic centroid connector placement is even more critical. Yet, very little literature exists exploring this topic. Jafari et al. explored two strategies for centroid connector placement in DTA and their impact on the resulting traffic flow (2015). The first methodology radially distributed access/egress points to the nodes nearest to the zonal centroid. This experiment brought to light some of the limitations to the methodology so commonly recommended in practice. The second methodology explored the use of "bi-level" assignment, which divided each TAZ into two concentric subzones and distributed the total TAZ demand between the two subzones according to an "inner-to-outer demand ratio", to better distribute demand throughout the TAZ. They found that this placement strategy produced results more consistent with real-world behavioral patterns (Jafari et al., 2015).

Two key aspects are evident from the review of available literature. Firstly, there has been limited investigation into centroid connector placement within STA and DTA models. Secondly, the studies which have been conducted indicate that centroid connector placement affects model outputs. Both these factors highlight the importance of the research presented in the thesis to provide an improved understanding of the impacts of centroid connector placement.

2.7 Application of Innovative Data Sources in Transportation

The International Transport Forum lists transportation operations, planning, and safety as three areas where big data has the potential to provide data-driven insights and transform transportation policy (Organisation for Economic Co-operation and Development, 2015). However, a brief review of literature indicates that most of the efforts, to date, have focused on big data applications for transportation operations. A report by the United States Department of Transportation Intelligent Transportation Systems Joint Program Office (USDOT ITS JPO) dedicated an entire chapter to how big data is currently being leveraged in transportations operations by both the private sector (e.g. Waze, MyRideBuddy, and AirSage) and the public sector (e.g. San Diego and Dallas Integrated Corridor Management programs and Michigan Department of Transportation's Data Use Analysis and Processing, Integrated Mobile Observations, and Weather Response Traffic Information programs) (Burt, Cuddy, & Razo, 2014). Additionally, in the summary of session highlights for the 2014 Transport for a Changing

World International Transport Forum Annual Summit, only applications of big data in transportation operations were highlighted. Lastly, recent research is mostly exploring the viability of leveraging big data to optimize active traffic management applications (Shi and Abdel-Aty, 2015; Yu, Park, Kim, & Ko, 2014).

Buckley and Lightman (2015) hypothesize that big data can be leveraged by transportation agencies in the transportation planning framework to develop more timely origin-destination matrices, better analyze route and mode choices, and create more robust network models. However, little research has been performed exploring the possibilities. Dong, Wu, Ding, Chu, Jia, and Qin (2015) explored the viability of using call detail records as a data-driven approach for traffic analysis zone division, trip generation, and trip distribution. Their approach offers unprecedented flexibility to select and divide the desired number of TAZs with reasonable accuracy and can be used by MPOs for the development of new travel demand forecasting models or improving existing models. Additionally, Toole, Colak, Strut, Alexander, Evsukoff, and Congalez (2015) used call detail records to generate origin-destination matrices and trip tables and found the generated trip tables were in close agreement to those generated by the 2011 Massachusetts Household Travel Survey in Boston and the 2000 Bay Area Travel Survey in San Francisco.

2.8 Parcel Data

Another important aspect of this project is the use of parcel data as a proxy for demand generation locations. Land parcel databases represent the physical location of residences, businesses, and public lands and form the basis for all land use and zoning decisions (National Research Council, 2007). Parcel data, or cadastral data, represent the most appropriate level of geographic detail for decisions related to the development of land, business activities, and emergency response (see Figure 2.4). Parcel data are also critical to urban planning and the analysis of transportation needs, environmental issues, and natural hazard risk (National Research Council, 2007).

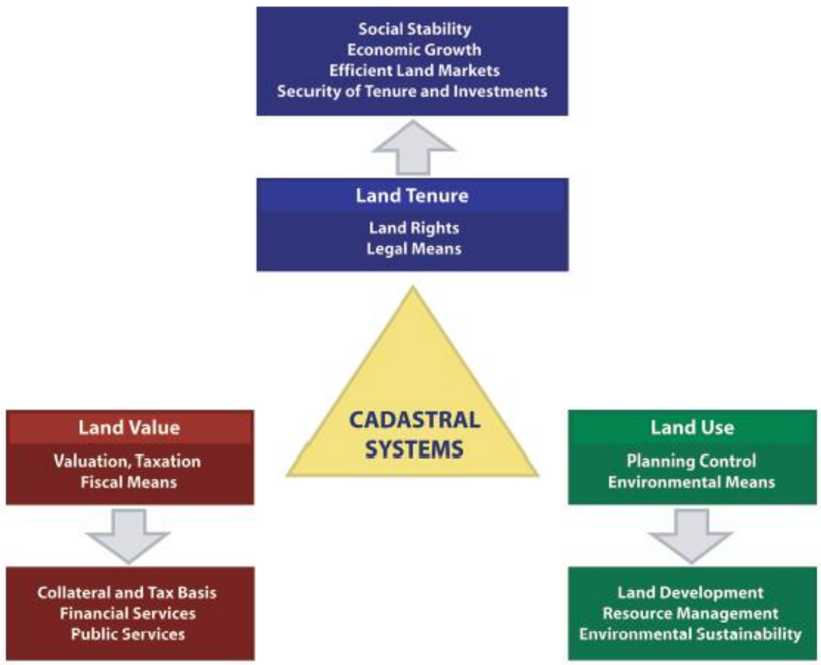


Figure 1.4: The importance of parcel data (National Research Council, 2007)

In 1980, a National Research Council (NRC) study, *Need for a Multipurpose Cadastre*, declared that parcel data should be the fundamental building block for a nationally integrated system of land information. However, such a system has not been created to date. As of 2007, 70% of tax parcels in the US are digitized; the remaining 30% are in the most rural counties (National Research Council, 2007). Almost 20 states have converted more than 80% of their parcel data to a digital format; West Virginia, South Carolina, and New Hampshire have the least with only approximately 10% of their data digitized. The challenge at hand is that the lack of a nationally integrated land parcel data framework, despite recommendations by the NRC, has led to duplicative efforts. Additionally, the fragmentation of land information has led to inconsistent quality and availability of land parcel information across US. The National Integrated Land System is the closest structure in place that resembles a coordinated effort for a national database, but, in its current condition, it is more of a set of technologies than a source of parcel data (National Research Council, 2007).

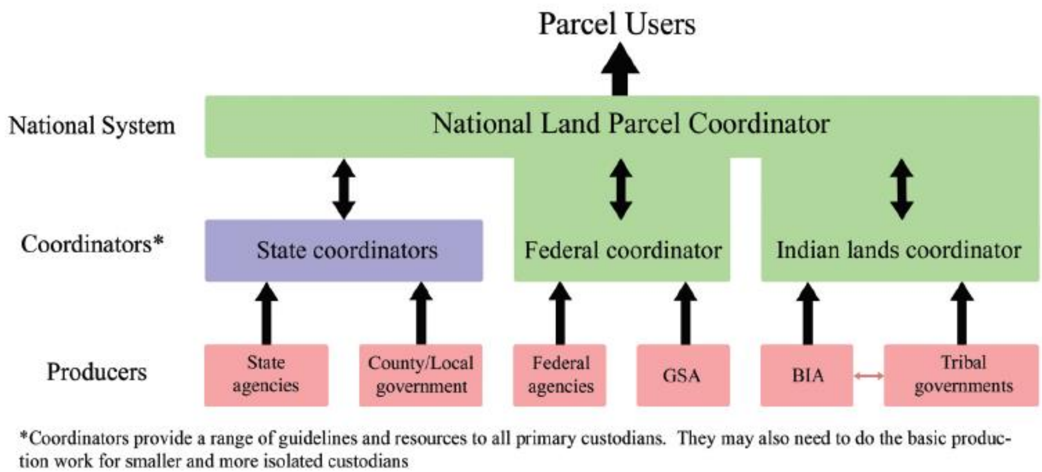


Figure 2.5: Flowchart of proposed national casdastre system (National Research Council, 2007)

The vision for a national cadastre system, recommended by the Committee on Land Parcel Databases is depicted in Figure 2.5. Local government would be the creator/maintainer of parcel data within a county or city. The state government would be responsible for the assembly of a comprehensive set of parcel data on an annual basis; it would also produce/maintain parcel data for counties who are not financially or technically able. Lastly, the federal government ties the system together by ensuring that federally managed land's parcel data is integrated into the system and that all other parcel data sources are integrated into the system properly (e.g. ensure state boundaries line up)

(National Research Council, 2007). If such a system comes to fruition, this would mean that the data necessary to employ this centroid connector placement methodology would be ubiquitously available for any major US city (see recommendations for data attributes in Figure 2.6).

Metadata—The metadata will contain information about the data such as the data steward, the parcel contact, the date of the data, and other information that would support the use and application of the information (e.g., projection, coordinate system).
Parcel Outline (Polygon)—The geographic extent of the parcel, the parcel boundaries forming a closed polygon.
Parcel Centroid (Point)—A point within the parcel that can be used to attach related information. This may be a visual centroid or a point within the parcel. It may not be the mathematical centroid because this point needs to be contained within the parcel polygon.
Parcel ID—A unique identifier for the parcel as defined by the data steward or data producer. The parcel identifier should provide a link to additional information about the parcel and should be unique across the data steward's geographic extent. (Note: To obtain a unique nationwide Parcel ID, the jurisdiction's Federal Information Processing Standards [FIPS] code could be added to the front of the parcel ID when the source data are posted to the publication site. This can be done transparently so that data producers do not have to worry about changing their parcel numbers. In the case of federally owned parcels, data concerning federal agency code and state FIPS code may be incorporated transparently as well.)
Owner Type—The classification of the owner. In some local governments, tax parcels are tagged as either taxable or exempt and the owner classification is not known. In these cases owner types of taxable and exempt may be added to this list: international, tribal, federal, state, county, local or municipal, private, not for profit, other, or unknown.
Parcel Street Address—The street address (site address) for the parcel. If there is more than one, select the first or primary site address. (Additional site address may be provided in a related file.)
Parcel Area—The area of the parcel expressed in acres.
Public Parcel Name—For publicly owned parcels, this is the commonly recognized name of the parcel (e.g., Dad Dunham Park or Yellowstone National Park).

Figure 2.6: Recommended attributes for national parcel data database (National Research Council, 2007)

In the literature, parcel data has been utilized as a way to approximate demand in several fields including water resources (Morales, 2010) and in earlier steps of the four step travel demand modeling process. Activity based models are beginning to explore the use of parcel-level data for demand modeling efforts. The Sacramento Area Council of Governments (SACOG) was the first major MPO to utilize parcel-level data in travel demand modeling efforts when they incorporated land-use policy and planning through its Preferred Blueprint Scenario in December 2004 (Griesenbeck, 2006). The SACOG activity based model (SACSIM) uses parcel-level data, instead of TAZs, in order to be able to answer questions about development patterns, street patterns, and proximity to transit services.

2.9 Conclusions

Though static traffic assignment is the predecessor decision support tool in transportation network analysis, it remains a tool that is very useful in certain situations, notably in academia and in practice where resources are limited or data are not known with high certainty. STA is highly robust to input data errors and provides a fast, transparent, stable, and provably correct equilibrium solution. However, due to limitations with the problem formulation, notably the steady state link performance functions, the resulting equilibrium flows are systematically overestimated on higher capacity corridors—with travel times on these corridors grossly underestimated due to link performance functions—and flow on lower capacity arterials and collectors is underestimated. Additionally, previous work has found that equilibrium solutions are not stable with respect to arbitrary centroid connector placement. One study in the literature produced more stable equilibrium flows by allocating centroid connectors according to an optimization algorithm that sought to minimize the volume-to-capacity ratio on characteristic links; however, achieving more realistic volume-to-capacity ratios does not necessarily make the results more representative of real world conditions.

Dynamic traffic assignment, another traffic assignment decision support tool that is capable of accurately capturing congestion propagation and queue spillback, is not presently being implemented on a widespread basis by metropolitan planning organizations. A survey by Cambridge Systematics of the 20 largest MPOs in the US (plus three MPOs known for their innovation) indicated that only two MPOs have DTA models for their entire metropolitan planning area (Cambridge Systematics, 2015). Seven others indicate that they have models under development, which are expected to be completed within the next five years. Two of the biggest barriers to wide-scale DTA deployment in MPOs are concerns related to acquiring the necessary input data and the time and effort required in calibrating and validating the model.

This thesis explores the use of readily available parcel data for a more data-driven approach to the allocation of centroid connectors and dispersion of TAZ demand in both static and dynamic traffic assignment. The goal is to advance prior work by developing more systematic guidance on how to accurately place centroid connectors. This work has two primary goals: reduce both the presently required visual inspection of connector locations and the necessary model calibration efforts. The requirement of manual visual inspection is minimized by creating a methodology that selects entry/exit locations consistent with reality and in locations that will not create artificial bottlenecks. Additionally, this methodology is evaluated for its ability to capture data-supported network flows. A review of available literature indicates that a process that achieves these objectives is largely absent.

Chapter 3: Methodology

3.1 Introduction

In this chapter, the framework to develop a centroid connector structure informed by parcel data, to divide demand throughout a TAZ by parcel density, and to integrate the new structures into existing traffic assignment paradigms is presented. In order to assess the impact of centroid connector placement on modeling results, static and dynamic traffic assignment models must be run to convergence; thus, this chapter begins by outlining the solution methodology for static and dynamic traffic assignment models identified in the literature. In addition, this chapter presents the logic for the new methodology that selects network nodes to act as network entry/exit points via centroid connectors and that allocates demand for a TAZ between two sub-zones based on the built environment. Finally, flow charts are presented showing the flow of work between the various software and Java/C codes in order to complete this analysis. The notation utilized in this chapter for static traffic assignment and the methodology for the development of centroid connectors and demand allocation by parcel data are presented in Tables 3.1 and 3.2 respectively.

Table 3.1: Notation Summary Static Traffic Assignment (Section 3.2)

Notation	Description
$G(N, A)$	network consisting of nodes (N) and arcs (A)
Z	set of zones
(r, s)	origin (r) destination (s) pair
x_{ij}	flow on link ij
t_{ij}	travel times on link ij as a function of the assigned flows x_{ij}
t_{ij}^{o}	free flow speed on link ij
c_{ij}	the "practical" capacity of link ij, set equal to the capacity of the link resulting in LOS C or 80% of the theoretical capacity
α and β	calibration factors for BPR function, often 0.15 and 4.00, respectively.
Π^{RS}	set of all acyclic paths connecting zone R to zone `S
h^{π}	number of travelers choosing path π
δ_{ij}^{π}	indicator variable (takes a value 1 if link ij lies on path π, 0 otherwise)
d^{rs}	number of trips between zone r and zone s
c^{π}	travel time on path π
$TSTT$	total system travel time
$SPTT$	"theoretical" total system travel time, if everyone were on their own shortest path

Table 3.2: Notation Summary for the Data-Driven Placement of Centroid Connectors and Allocation of Demand by Parcel Data (Section 3.4)

Notation	Description
Z_i	set of nodes that belong to the i^{th} TAZ
I_i	set of nodes that belong to the inner zone of the i^{th} TAZ
r_i	threshold value for zone i, dividing the zone into its inner and outer subzone
N_i	number of nodes in zone i
d_j^i	Euclidean distance from zone i's centroid to node j
w_i	weighted portion of demand falling spatially within the inner subzone
w_o	weighted portion of demand falling spatially within the outer subzone
n_i	parcel weight assigned to node i
d_i	the demand assigned to the network via connectors in inner subzone
d_o	the demand assigned to the network via connectors in outer subzone
D	total demand for the TAZ

3.2 Static Traffic Assignment

In STA, the fundamental behavioral assumption of traffic assignment, the principle of user equilibrium, follows Wardrop's first principle: for all origin-destination (O-D) pairs, all used paths will have equal and minimal travel times (Wardrop, 1952). The static assignment problem can be formulated as an optimization model, Equations 2 through 5, seeking to minimize the Beckmann function (Equation 2) (Beckmann, McGuire, & Winsten, 1956).

$$\min_{x,h} \sum_{i,j \in A} \int_0^{x_{ij}} t_{ij}(x)dx \tag{2}$$

$$x_{ij} = \sum_{\pi \in \Pi^{RS}} \delta_{ij}^{\pi} * h^{\pi} \; \forall \, i,j \; \in A \tag{3}$$

$$d^{RS} = \sum_{\pi \in \Pi^{RS}} h^{\pi} \; \forall \, (r,s) \in Z^2 \tag{4}$$

$$h^{\pi} \geq 0 \; \forall \, \pi \; \in \; \Pi^{RS} \tag{5}$$

This mathematical program considers a networking comprising of N nodes, A arcs and Z travel zones. The first constraint, Equation 3, requires that the flow on each link ij (x_{ij}) must be exactly equal to the total flow for all paths connecting origin-destination (O-D) pair rs if link ij lies on the shortest path π connecting rs. The second constraint maintains that all demand for each O-D pair must be satisfied and assigned to a path π, conserving the demand between origin and destination pairs across the network (Equation 4). Non-negativity is maintained through constraint 3 (Equation 5) ensuring that no path flows are negative.

Though there's not an intuitive meaning of the objective function, it was derived through back calculation, by Beckmann et al. (1956), from the following optimality conditions:

$$h^{\pi} \geq 0 \; \forall \, \pi \; \in \; \Pi^{RS} \tag{6}$$

$$c^{\pi} \geq \kappa^{rs} \;\; \forall \, (r,s) \; \in Z^2 \tag{7}$$

$$h^{\pi}(c^{\pi} - \kappa^{rs}) = 0 \; \forall \, \pi \; \in \; \Pi^{RS} \tag{8}$$

Optimality condition 1 ensures that all path flows are nonnegative (Equation 6). Optimality condition 2 shows that κ^{rs} is the shortest path travel time between O-D pair rs (Equation 7). Optimality condition 3 shows that if a path π is used, the cost to traverse the path, c^{π}, must be equal to the shortest path travel time, κ^{rs}; conversely, if the travel time on path π (c^{π}) is longer than the shortest path travel time for the O-D pair rs (κ^{rs}), the flow on path π (h^{π}) must be 0 (Equation 8). The novelty and elegance of this static traffic assignment formulation is that the gradient of the objective function is the shortest path problem, allowing for gradient based optimization methods to determine a user equilibrium solution.

The framework to obtain such a solution is iterative, starting with an initialized set of path flows and working toward user equilibrium. An example pseudocode can be seen below:

1. Determine a feasible link flow solution (x, h). Initialization of the network is normally completed via all or nothing assignment (assign all vehicles to the shortest path with travel times calculated assuming zero flow on the network).
2. Calculate the link travel times via the selected link performance function using the flows x.

3. Find the shortest path between all origins and destinations. This can be completed using a label setting algorithm (e.g. Dijkstra's algorithm) or a label-correcting algorithm (e.g. Bellman-Ford algorithm).
4. Assign all of the demand between each origin-destination pair to the shortest path. Let this be h^*. Let x^*be the link flows corresponding to h^*.
5. Shift some of the travelers onto the newly identified shortest path between each O-D pair. Update the link flow via the following equation:

$$x \leftarrow \lambda x^* + (1 - \lambda)x \qquad (9)$$

By shifting the flows in an incremental fashion, it avoids oscillation between solutions. Thus, the selection of λ is critical. Two options for finding λ identified in the traffic assignment problem (TAP) literature include the Method of Successive Averages (MSA) or Frank-Wolfe algorithm.
6. If the updated solution satisfies a pre-defined convergence criterion, then the pseudocode terminates. Else, return to step 2. The convergence criteria measures how close the solution is to equilibrium conditions; this concept has been defined in a number of ways. One of the more accepted termination criteria for shortest path assignment is when the average excess cost drops below a predefined threshold. Average excess cost, which is normalized to show how much longer the average vehicle trip is compared to the shortest path travel time, is calculated by the equation shown below:

$$AEC = \frac{TSTT - SPTT}{\sum_{rs} d^{rs}} \qquad (10)$$

3.3 DYNAMIC TRAFFIC ASSIGNMENT

Dynamic traffic assignment (DTA) models account for temporal and spatial dependencies of travel behavior and the subsequent impacts on traffic flow across a network. As mentioned within the literature review, DTA methodologies can avoid the limitations of static modeling, realistically accounting for the development of congestion (e.g. queuing, spillback, lane changing and merging behavior) and real-time information based impacts on route choice (Balakrishna et al., 2013, Chiu et al., 2011). Observing the impact of centroid connector placement from a dynamic context is vital from a demand, routing, and capacity assessment perspective.

As mentioned in Section 2.3, there are a number of variations for which DTA formulations have been created. In general, most DTA models are based on the extension of the static user equilibrium principle to create "dynamic user equilibrium" (DUE), which is a time-dependent version of the Wardropian equilibrium: all used paths have equal and minimal travel times for each origin-destination pair *and* departure time.

The concept of DUE can be presented considering discretized time (Merchant and Nemhauser 1978a, Merchant and Nemhauser 1978b) and continuous time (Boyce et al., 2001, Friesz et al., 1989, Friesz et al., 1993). The modeling conducted within this thesis utilizes VISTA, a simulation-based platform approach using a cell transmission model, considers discretized time and is described in Section 2.4. Accordingly, the DUE formulation applied within the software originates from the seminal work of Merchant and Nemhauser in 1978.

One of the more straight forward steps in static traffic assignment is the calculation of link travel times using link performance functions and then adding the links associated to specific paths to determine path travel times. In contrast, this is the most complex step in DTA and is known as the network loading problem (NLP). The

NLP uses link models and node models to simulate traffic flow throughout the network based on fundamental diagrams (FD) of traffic flow. Link models, summarized in Figure 3.1, model what would happen on a single roadway link and yield the sending (downstream) and receiving (upstream) flows based on the Lighthill-Whitham-Richards (LWR) model of traffic flow as a fluid (Lighthill and Whitham, 1955; Richards, 1956). Node models (see Figure 3.2) utilize the sending flow of the upstream link and the receiving flow of the downstream link to find the number of vehicles physically allowed to travel from one link to the next based on prevailing congestion conditions. The list of node models in Figure 3.2 is far from exhaustive, as models also exist for control devices at intersections (e.g. signals, yield signs, and roundabouts) as well as those more concerned with capturing the behavior of drivers with turn taking or gap acceptance (Tampere, Corthout, Cattrysse, & Immers, 2011; Corthour, Flotterod, Viti, and Tampere, 2012).

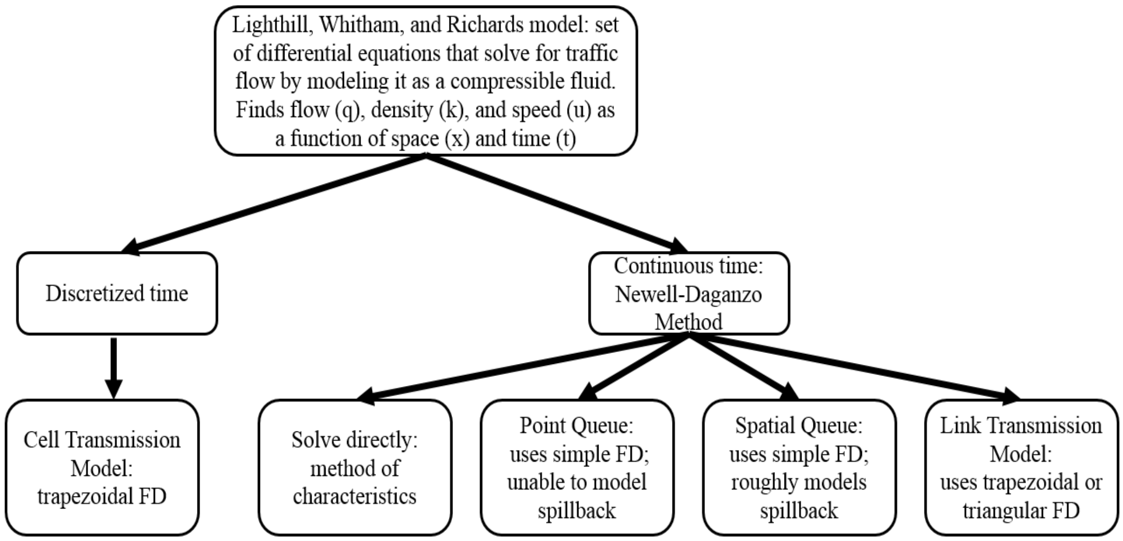


Figure 3.1: Examples of Link Models commonly used in the NLP

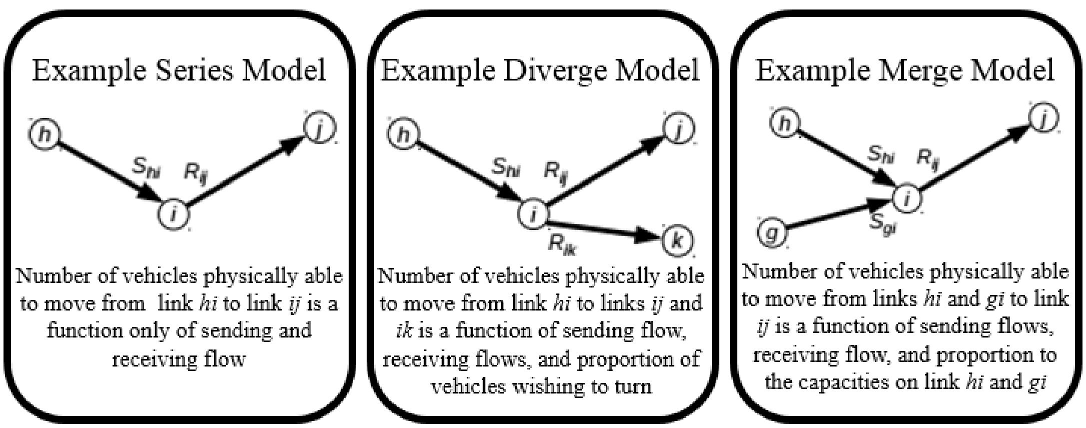


Figure 3.2: Examples of Node Models in the Literature

Upon completion of the network loading problem, the DTA methodology closely follows that of static traffic assignment. Thus, a potential pseudocode is as follows:

1. Start with initial path flows, H_0. This can be found by placing all or a fraction of travelers on their shortest path assuming zero flow (like the first step in STA), placing travelers according to the optimal solution found with static assignment, or some other "warm start" methodology (Nezamuddin, 2011).
2. Find the travel times on each path using the network loading procedure described above. This yields a travel time matrix, T, which reports the travel time on each path for each departure time.
3. Find the shortest path between each O-D pair at each departure time after network loading using a time-dependent shortest path (TDSP) algorithm (Ziliaskopoulos and Mahmassani, 1993).
4. Subsequently, the target path flow matrix, H^*, is obtained by placing all travelers on the available shortest path at their departure time.
5. Update the solution for the path flows matrix by taking a weighted average of the path flow matrix, H, and the target path flow matrix based on TDSP after network loading, H^*. This can be completed via a convex combination of the two solutions (e.g. using the method of successive averages to find the step size, λ), simplicial decomposition, or gradient projection.
6. If the updated solution satisfies a pre-defined convergence criteria, then the pseudocode terminates. Else, return to step 2. The convergence criteria measures how close the solution is to equilibrium conditions. One common termination criteria is the average excess cost, or the average cost

of all paths for all departure times, weighted by the total number of people on each path at each departure time. In dynamic traffic assignment, AEC is defined as:

$$AEC = \frac{\sum_{\pi,t} h_{\pi,t}\,\tau_{\pi,t} - \sum_{\pi,t} h^{*}_{\pi,t}\,\tau_{\pi,t}}{\sum_{\pi,t} h_{\pi,t}} \tag{11}$$

3.4 Algorithm for Connector Placement via Parcel Density

The focus of this research is to develop an automated centroid connector placement strategy that uses parcel data to generate network entry/exit points consistent with the likely location of activities within a TAZ, thus achieving more realistic traffic patterns on local and major streets.

This research builds on the work of Jafari et al. (2015), who found that the distribution of centroid connection points throughout a TAZ—as opposed to either the locations nearest the zonal centroid or along the zonal boundary, both of which are suggested in the literature—achieves more realistic traffic patterns on networks with detailed representation of lower functional class roadways (2015). The predecessor work utilized a network-wide, user-defined demand split between two concentric zones dividing the TAZ and a radial distribution of a user-defined number of centroid connectors to eligible nodes nearest the zonal centroid in each subzone. This work differentiates itself by utilizing built environment parcel data to automate the selection of TAZ-specific inner-to-outer subzone demand splits and to select a user-defined number of areas of high development as the appropriate locations to connect TAZ centroids to the network in each subzone of a TAZ.

The approach is straightforward. A list of eligible nodes is created by filtering out nodes that are undesirable network entry/exit locations (e.g. nodes on a limited access facility or at a signalized intersection). The residential and commercial land use parcels are then assigned to the nearest eligible node using a geospatial analysis tool in order to "weight" the nodes, or differentiate the nodes that were more likely to be actual entry/exit points in the real world. Additionally, the Euclidean distance of each eligible node to the centroid of its TAZ is computed. This information is critical in order to create a zone-specific "threshold" for each TAZ; this threshold is used to split the zones into a concentric inner and outer subzone, found in the literature to achieve more realistic network loading patterns. The demand split between the two concentric subzones in each TAZ is determined uniquely using the geographic dispersion of parcel density in each TAZ. The methodology then selects the highest n weighted nodes in each subzone of a TAZ to serve as the new connector nodes. The following sections provide further detail on each component of the proposed approach, including the initial data processing steps. The methodology was implemented in Java, while most of the pre-processing steps were accomplished using GIS software. The only user inputs required for this methodology are the determination of eligible nodes and the number of connectors per subzone. A summary of the notation used in this section can be found in Table 3.2. A pictorial representation of this placement of centroid connectors via parcel data can be found in Figure 3.3.

3.4.1 Data Preprocessing

Data preprocessing was necessary for both parcel-level and traffic network data. For the latter, the proposed methodology requires distinguishing between centroid nodes and regular "eligible" nodes. The "eligible" list only includes the network nodes that are reasonable entry/exit points (e.g. excludes nodes on limited access facilities). GIS software was used to assign a TAZ ID to each eligible node through a one-to-many approach. This allows nodes along TAZ borders to be assigned to each neighboring zone, thus making them eligible connection points for any of the corresponding centroids. As such, the match operation selected in the software tool assigns nodes a particular TAZ ID if it was within a specified distance of the TAZ boundary. For this particular network, a buffer distance of five feet led to the inclusion of nodes that lay along the border between TAZs without adding superfluous nodes.

Parcels, originally geocoded as polygons, were translated into point data using GIS software to facilitate the algorithmic implementation of the proposed methodology. Relevant information for each parcel from the original shapefile includes coordinate data, built square footage, and classification of the land use of the parcel (e.g. residential, commercial, etc.). The necessary information to obtain from GIS for the new methodology includes the TAZ that a parcel falls within and the nearest eligible network node. This was completed using the spatial analysis tool. The built square footage of every residential and commercial land parcel was assigned to only one TAZ, as most TAZ boundaries are physical barriers (e.g. rivers, major roads, etc.), and to the node nearest to the geometric centroid of the parcel by Euclidean distance. Each parcel was assigned to a single node and TAZ to avoid inappropriately biasing the demand distribution. This seeks to approximate reality, as the most densely developed areas in a TAZ are presumably going to generate the most demand on the real-world network.

3.4.2 Dividing the TAZ into an Inner and Outer Subzone

The algorithm for the data-driven placement of centroid connectors first divides the TAZ into two concentric areas: an inner subzone and an outer subzone (see Figure

3.3b). The motivation behind this decision was supported by previous research: having two subzones was found to achieve a more even distribution of entry/exit points throughout the TAZ and encourages the simulated vehicles to use local streets in a manner consistent with real world behavior (Jafari et al., 2015). The inner subzone's radius was computed as the average distance between each node in the selected TAZ (i.e. each potential entry location onto the network) and the TAZ's centroid, as shown in Equation 12:

$$r_i = \frac{1}{N_i} \sum_{j \in Z_i} d_j^i \tag{12}$$

where r_i is the threshold value for zone i, N_i is the number of nodes (potential entry locations) in zone i, and d_j^i is the Euclidean distance from zone i's centroid to node j. Z_i defines the set of nodes that belong to zone i. In terms of implementation, the process effectively splits each centroid into two: a sub-centroid for the inner subzone and a sub-centroid for the outer subzone. This allows the demand split between the two subzones to be enforced.

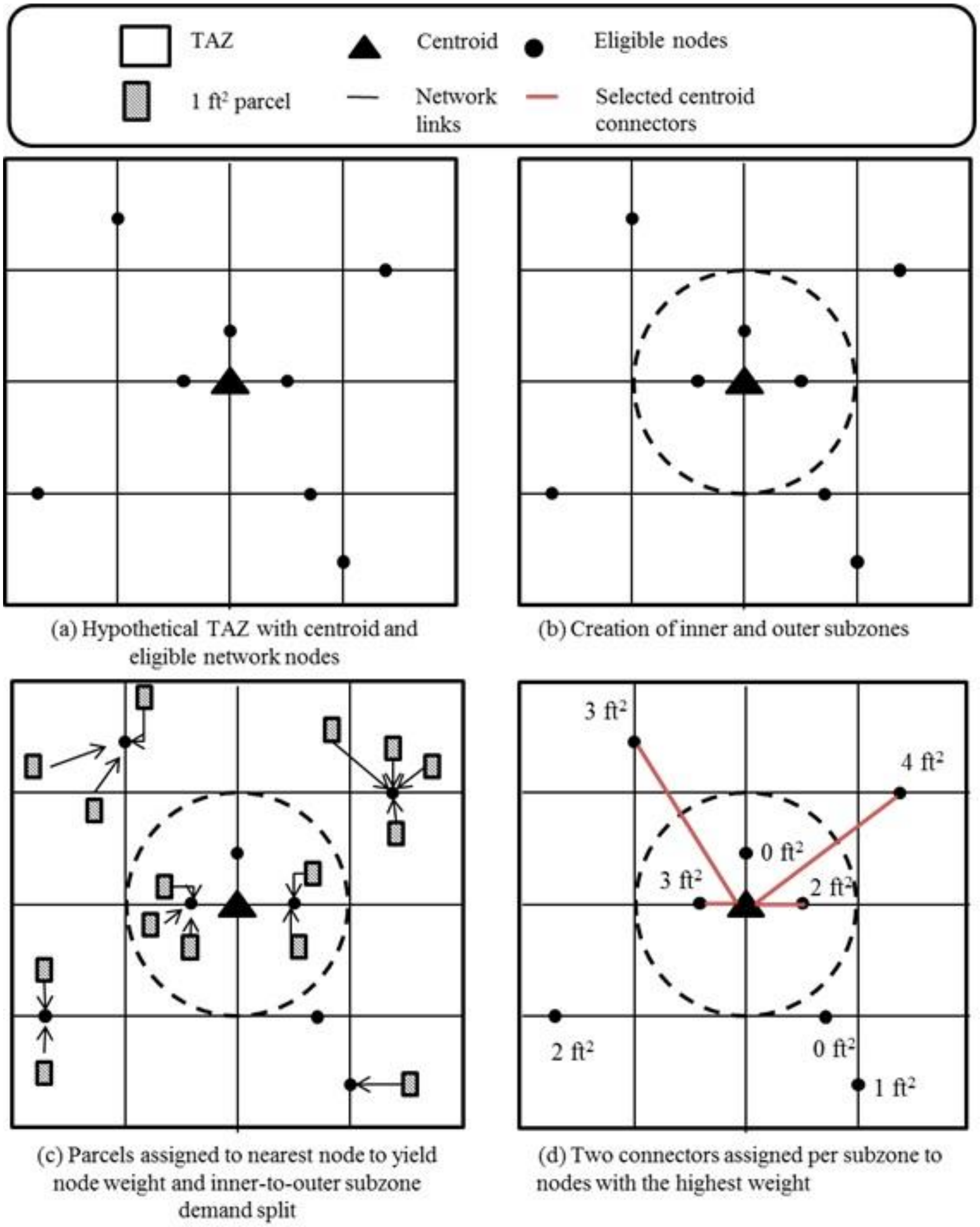


Figure 3.3: Pictorial Representation of Methodology

3.4.3 Determining the Demand Split for the Subzones

Based on the data produced by the preprocessing described in Section 3.4.2, each parcel was assigned to the nearest eligible network node by Euclidian distance (see Figure 3.3c); note that a node n is considered accessible to a zone z if a parcel belonging to zone z is assigned to node n. Parcels, especially at zone z's boundary, may be assigned to a node n in a different zone z', and in this way that node n is accessible from zone z' even though node n is outside the zone z''s boundary; this is argued as an acceptable process as, in reality, zone boundaries do not exist.

The weight of each node n is equal to the total built square footage of parcels assigned to it. The demand split for each TAZ's inner and outer subzone was computed by summing the weight (built square footage) of all nodes that fall spatially within each subzone and dividing by the total weight (built square footage) of all parcels that are accessible to the TAZ. This is demonstrated mathematically in Equations 13 and 14:

$$w_i = \frac{\sum_{i \in I_i} n_i}{\sum_{z \in Z_i} n_z} \tag{13}$$

$$w_o = 1 - w_i \tag{14}$$

where w_i is the weighted proportion of the demand that belongs to the inner subzone, n_i is the parcel weight assigned to node i, n_z is the parcel weight assigned to node z, w_o is the weighted proportion of the demand that belongs to the outer subzone, I_i is the set of nodes that belong to the inner zone of the i^{th} TAZ, and Z_i is the set of nodes that belong to the i^{th} TAZ where I is a subset of Z.

For example, if a TAZ housed parcels that summed to 100,000 square feet, where 75,000 square feet were assigned to nodes that fell spatially within the inner subzone, the demand ratio would be 3-to-1 inner-to-outer split. Thus, if the representative centroid for this TAZ had a demand of 100 vehicles, 75 vehicles would be assigned to enter the network within the inner subzone and the remaining 25 vehicles would enter through nodes in the outer subzone. This is explained mathematically in Equations 15 and 16:

$$d_i = w_i * D \tag{15}$$

$$d_o = w_o * D \tag{16}$$

where d_i is the demand that is assigned to the network via connectors in the inner subzone, d_o is the demand that is assigned to the network via connectors in the outer subzone, and D is the total demand for the TAZ.

This represents an advancement of prior work, as the demand split is unique to each TAZ and not a “one-size-fits-all” approach. There are two exceptions to this rule: in the case of a subzone having no parcel information, or zero built square footage, 100 percent of the demand is assigned to the other subzone; in the case of no parcels residing in a TAZ, or zero built square footage, the demand is split 50/50 between the two subzones. The split of 50/50 was selected because it was argued that without additional network, demand, or development detail in the area, a rational case cannot be made to allocate demand asymmetrically.

3.4.4 Selecting the Entry/Exit Nodes

The number of centroid connectors (n) per subzone is a user input variable requiring sensitivity analysis, which is consistent with prior studies. This project builds on the methodology identified in prior research for selecting entry/exit nodes. In work by Jafari et al. (2015), n entry nodes are selected based on their distance from the TAZ centroid. In this methodology, the entry nodes for each subzone are the n highest weighted nodes in each subzone (see Figure 3.3d). In the case of a TAZ with no parcels, the entry nodes are selected using the previous method whereby the n nodes chosen in each subzone are the closest, defined by Euclidian distance, to the centroid of the parcel-less TAZ.

3.5 Implementation

The work flow for this project is shown in Figures 3.4 and 3.5 for static traffic assignment and dynamic traffic assignment, respectively. The input data manipulation is completed using ArcGIS, though the results are replicable with any spatial analysis tool. The creation of the new centroid connector structure and demand profile, updated using the inner-to-outer demand ratio, is completed via a Java code, created by Ehsan Jafari, and run directly from the file server, where the model and the code are stored. The resultant network structure and demand profile are uploaded directly into VISTA for analysis (DTA) or export (STA). The static assignment code utilized in this analysis was provided by Dr. Stephen Boyles and is available online at https://tinyurl.com/SteveBoyles under CE 392C: Transportation Network Analysis 1. Additionally, Michael Levin provided a base code to aid in the data manipulation to export the network from VISTA into a format compatible with the static assignment code. The selected DTA software is VISTA (Section 2.4).

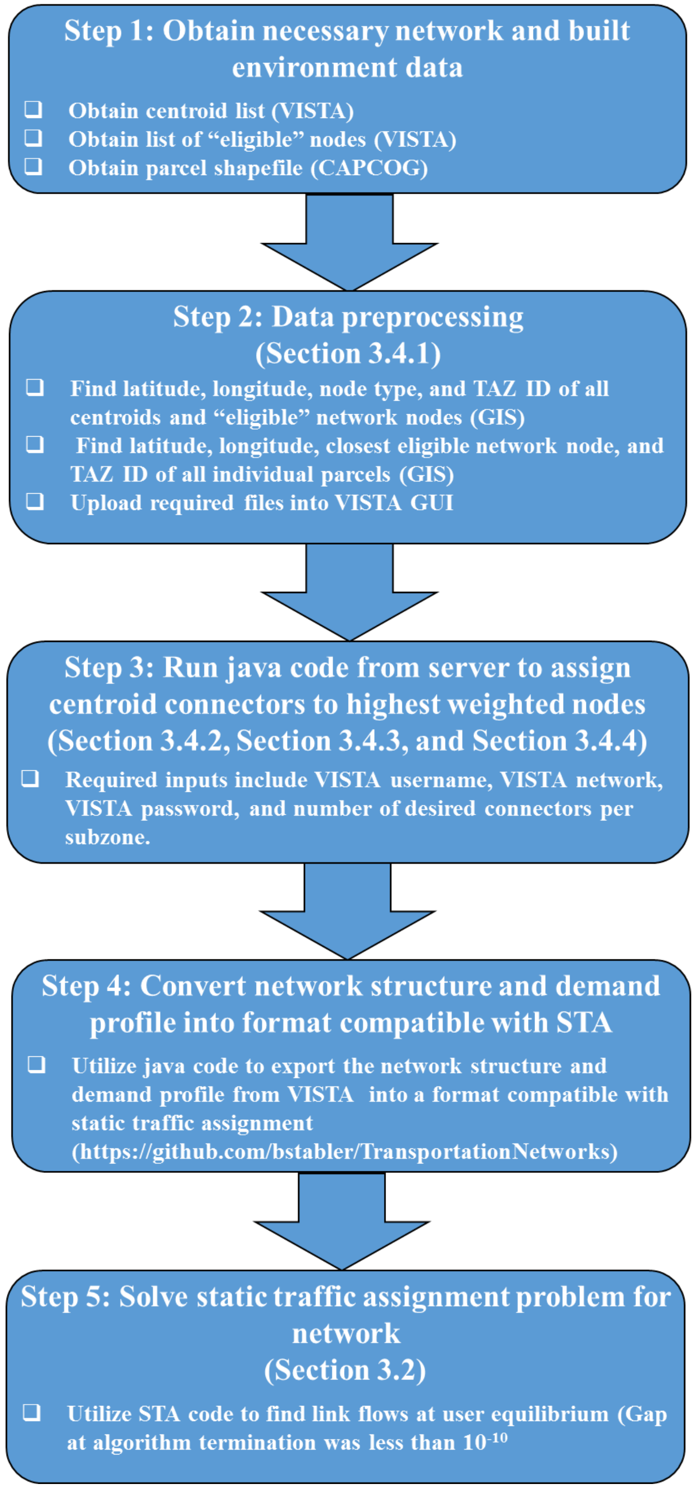


Figure 3.4: Static Traffic Assignment Workflow

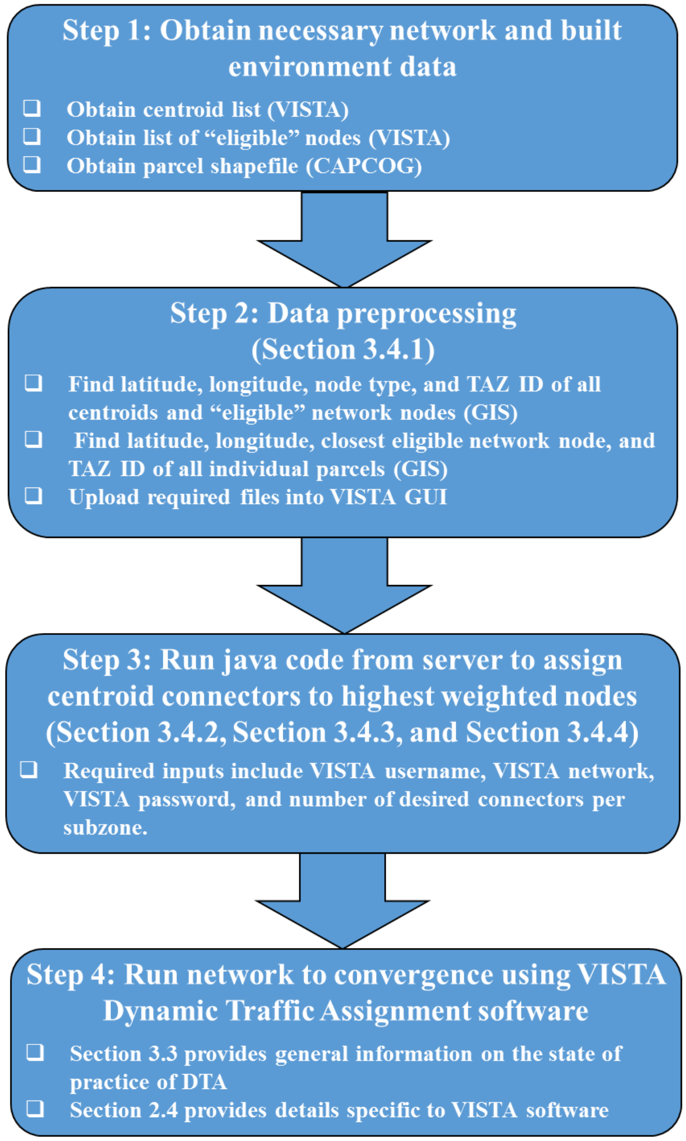


Figure 3.5: Dynamic Traffic Assignment Workflow

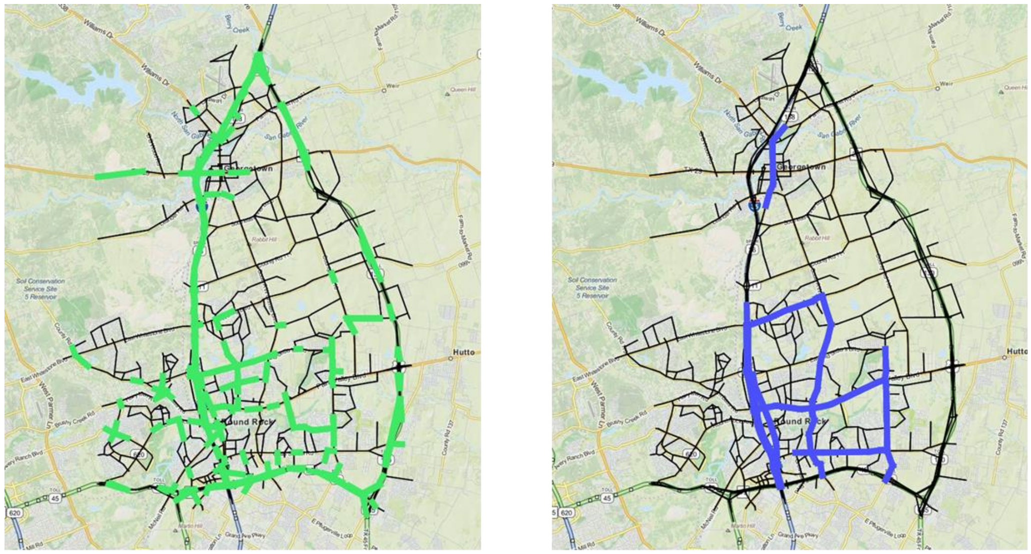
Chapter 4: Network Data and Experimental Design

4.1 Introduction

This section describes the numerical experiments conducted to assess the performance of the methodology described in the previous chapter. The goal of these experiments is to explore whether using parcel data allows for a better approximation of vehicle entry/exit points, ultimately leading to more accurate models. Field counts and travel times along major corridors are used to assess model performance.

4.2 Network Description

Experiments were conducted using a sub-area network in the Austin, TX region located in Williamson County. The network topology and attributes were extracted from the Capital Area Metropolitan Planning Organization's (CAMPO) regional model, and refined to incorporate additional roadway detail throughout the network. The resulting base network includes 3,440 links, 1,680 nodes, 399 centroids, and 823 centroid connectors and supports a demand of 135,616 vehicles. Traffic signal data were provided by state and local agencies and entered into the model. Sub-area demand for the AM peak period (6 a.m. - 9 a.m.) was extracted from a regional DTA model. Available field data collected to calibrate the network includes counts on 1,305 network links and travel times along 18 corridors (Figure 4.1). Parcel data were obtained from GIS files provided by the Capital Area Council of Governments (CAPCOG).



(a) Links with field count data available

(b) Corridors with field travel time data available

Figure 4.1: Network and Validation Data

4.3 Scenario Description

A total of fifteen different scenarios were modeled in this effort—seven in DTA and eight in STA, with five scenarios appearing in both analyses. The selected DTA software package, VISTA, is simulation-based and finds equilibrium solutions expected to represent recurrent congestion patterns (Waller & Ziliaskopoulous, 2000). The static assignment code utilized in this analysis was provided by Dr. Stephen Boyles and available online at https://tinyurl.com/SteveBoyles under CE 392C: Transportation Network Analysis 1.

The base scenario for both analyses includes the original centroid connector structure provided in the CAMPO regional model developed using the TransCAD software package.

4.3.1 Summary of DTA Scenarios

For comparison purposes, two scenarios were built using the aforementioned bi-level approach developed by Jafari et al. (2015), the critical predecessor to this work. Two inner-to-outer demand split ratios, 50/50 and 90/10, were considered for further implementation and testing of the bi-level method as part of this research effort based on their performance in previous research. The 90/10 inner-to-outer subzone demand split was selected because it was found to have the highest locality factor; the 50/50 inner-to-outer subzone demand split was selected because it was the ratio that marked a threshold of stability in the network total system travel time (Jafari et al., 2015).

In order to test the data-driven placement of centroid connectors presented in this study, four additional scenarios were developed using parcel data to determine the demand split between subzones: two with two connectors per subzone and two with four connectors per subzone. Much of the available literature suggests that using signalized intersections as entry/exit points for centroids should be avoided as it may create artificial congestion at these locations. Thus, alternate scenarios for the two connector and four connector case studies were implemented where nodes at signalized intersections were eliminated from eligible entry points in the network. A summary of these scenarios is presented in Table 4.1.

Table 4.1: Summary of DTA scenarios

Scenario	Description
Base	• Base network exported from CAMPO's regional model
Bilevel, 50/50 Demand Split	• Jafari et al., 2015 • Created two subzones for each TAZ and divided demand 50/50 between the subzones for all TAZs
Bilevel, 90/10 Demand Split	• Jafari et al., 2015 • Created two subzones for each TAZ and divided demand 90/10 between the subzones for all TAZs
Parcel Based, 2 Connectors, Signals Permitted	• Demand divided into two subzones with demand allocation determined by parcel density • 2 connectors placed at highest weighted nodes in each subzone of a TAZ • Connection points were allowed at signalized intersections
Parcel Based, 2 Connectors, Signals not Permitted	• Demand divided into two subzones with demand allocation determined by parcel density • 2 connectors placed at highest weighted nodes in each subzone of a TAZ • Connection points were not allowed at signalized intersections
Parcel Based, 4 Connectors, Signals Permitted	• Demand divided into two subzones with demand allocation determined by parcel density • 4 connectors placed at highest weighted nodes in each subzone of a TAZ • Connection points were allowed at signalized intersections
Parcel Based, 4 Connectors, Signals not Permitted	• Demand divided into two subzones with demand allocation determined by parcel density • 4 connectors placed at highest weighted nodes in each subzone of a TAZ • Connection points were not allowed at signalized intersections

4.3.2 Summary of STA Scenarios

The two and four connector strategies using parcel data and the base scenarios were utilized identically in the static traffic assignment analysis, with no necessary alterations of the network created for analysis in VISTA. Three additional scenarios were created independently from the DTA analysis. Based on the literature, there was concern that simulated vehicles would abide by the demand split by subzone, but choose to utilize the connector in the subzone that places them the closest to their destination or to a high-speed facility, given the inability to model queue spillback and congestion propagation (Qian and Zhang, 2012). Thus, two new scenarios were created using one connector per subzone, one with connection at signalized intersections allowed and the other without. These were created to force simulated vehicles to enter the network at the location mostly likely to be a high demand generator in each subzone, based on built environment data alone, thus limiting the ability of the demand to load simply where convenient.

However, the literature also indicates that too few connectors can create artificial congestion in STA at entry points. Thus, this same logic was carried out with two connectors per subzone; this required splitting each TAZ centroid in the network into eight components—one sub-centroid per connector, with two subzones per TAZ and two connectors per subzone. For example, for a given TAZ with a demand of 100, let's assume that 80% of the parcel density falls within the inner subzone. Per Section 3.4.3, 80 vehicles will be assigned to the inner subzone and 20 vehicles will be assigned to the outer subzone. Now let's assume that the two highest weighted nodes in the inner subzone both have 2,000 square feet assigned to each entry/exit node. In the prior scenarios, the 80 vehicles are free to choose which of the two connectors they wish to utilize based on the shortest path algorithm, not necessarily respecting the parcel density. Thus, the "micromanaged" scenario was created to ensure the demand split at the node level was consistent with the parcel density by creating a sub-centroid for each connector and assigning the properly portioned demand to each sub-centroid. Thus, in the above scenario, 40 vehicles will enter at one node in the inner subzone of the TAZ, while the remaining 40 vehicles enter at the other node, regardless of which one would provide a more “convenient” route. This greatly increases the size of the origin-destination matrix and is not realistic in large DTA networks, but was worthy of analysis in STA due to the model’s efficiency gains elsewhere. A summary of these scenarios is presented in Table 4.2.

Table 4.2: Summary of Scenarios for STA analysis

Scenario	Description
Base	• See Table 4.1 for description
Parcel Based, 1 Connector, Signals Permitted	• Demand divided into two subzones with demand allocation determined by parcel density • 1 connector placed at highest weighted node in each subzone of a TAZ • Connection points were allowed at signalized intersections
Parcel Based, 1 Connector, Signals not Permitted	• Demand divided into two subzones with demand allocation determined by parcel density • 1 connector placed at highest weighted node in each subzone of a TAZ • Connection points were not allowed at signalized intersections
Parcel Based, 2 Connectors, Signals Permitted	• See Table 4.1 for description
Parcel Based, 2 Connectors, Signals not Permitted, "Micromanaged" demand	• Demand divided into two subzones with demand allocation determined by parcel density • 2 connector placed at highest weighted node in each subzone of a TAZ • Demand entered/exited network proportionally to parcel density at entry/exit node through creation of sub-centroid for each connector • Connection points were not allowed at signalized intersections
Parcel Based, 2 Connectors, Signals not Permitted	• See Table 4.1 for description
Parcel Based, 4 Connectors, Signals Permitted	• See Table 4.1 for description
Parcel Based, 4 Connectors, Signals not Permitted	• See Table 4.1 for description

Chapter 5: Results

5.1 Introduction

This section describes the results from the numerical analyses in terms of the updated centroid connector structure and the corresponding model performance. For convenience, a summary of the 15 scenarios analyzed can be found in Table 5.1; for additional detail concerning the motivation for why each scenario was created, please see in Tables 4.1 and 4.2 for static and dynamic traffic assignment, respectively. All scenarios are analyzed with respect to the base model from CAMPO in order to provide insight on how this methodology improves the state of practice.

The static assignment results are presented first. These results are analyzed with respect to the methodology's ability to output "behaviorally consistent" results (e.g. output model results that increase the flow on lower capacity links and decrease the flow on higher capacity links), which are performance metrics suggested by Friedrich and Galster (2009) and Qian and Zhang (2012). Part of the novelty of this research is that available field traffic counts as described in Section 4.2, are used to assess if the more "behaviorally consistent" results are better capturing real world behavior (flows); other research in the literature uses stability of results or the ability to capture realistic route choice behavior, instead of consistency with field data, as a performance metric in STA.

Next, the dynamic traffic assignment results are presented. DTA results are analyzed with respect to the methodology's ability to accurately represent loading consistent with the built environment of the subnetwork and for the ability of the methodology to produce more realistic travel times and link counts without manual refinement of the connector structure or excessive calibration efforts. Visual inspection of the network, the average parcel density per entry/exit location, and the number of entry/exit points with zero square footage assigned are analyzed to assess the ability of the methodology to capture nature of the built environment within the subnetwork. Field corridor travel times are compared to the resultant travel times on the corresponding corridor in the converged model to assess the ability of the methodology to better capture reality; model link flows are analyzed in the same manner to assess the ability of the model to replicate realistic traffic counts.

Table 5.1: Summary Table of Scenarios

Static Traffic Assignment	Dynamic Traffic Assignment
Base	Base
-----	Jafari et al, 2015 Bilevel 50/50 Demand Split
-----	Jafari et al, 2015 Bilevel 90/10 Demand Split
Parcel Based 1 Connector Signals not Permitted	-----
Parcel Based 1 Connector Signals Permitted	------
Parcel Based 2 Connectors Signals not Permitted	Parcel Based 2 Connectors Signals not Permitted
Parcel Based 2 Connectors Signals not Permitted "Micromanaged" demand	-----
Parcel Based 2 Connectors Signals Permitted	Parcel Based 2 Connectors Signals Permitted
Parcel Based 4 Connectors Signals not Permitted	Parcel Based 4 Connectors Signals not Permitted
Parcel Based 4 Connectors Signals Permitted	Parcel Based 4 Connectors Signals Permitted

5.2 Parcel Methodology's Impact on Static Assignment Results

In the literature where centroid connectors and their impact on the results of traffic assignment models are of concern, methodologies are evaluated in one of two ways: either (a) by their ability to show that one methodology better achieves more realistic volume-to-capacity (V/C) ratios compared to another or (b) that there's a change in resultant flows through the network that causes them to be more behaviorally consistent as a result of the new connector structure (Friedrich & Galster, 2009; Qian & Zhang, 2012). Table 5.2 summarizes how the resultant flow on functionally classified links changes as a direct result of the altered centroid connector structures, summarized in Section 4.3.2. The percent change in flow between scenarios is calculated by Equation 17:

$$Percent\ change = \frac{Flow_i - Flow_{base}}{Flow_{base}} * 100 \tag{17}$$

where $Flow_i$ is the resultant flow on each link of the specified functional classification in the seven new scenarios and $Flow_{base}$ is the resultant flow in the base scenario. Thus, for example, the parcel-based methodology with one connector per subzone per TAZ and connector placement not permitted at signalized intersections increased flow on local links by 93 percent and decreased flows on principal arterials by 8 percent compared to the original base centroid connector structure provided by CAMPO.

Table 5.2: Percentage Change in Flow between Scenarios

	Local	Collector	Minor Arterial	Principal Arterials
Base	Base network is reference case			
Parcel Based, 1 Connector, Signals not Permitted	93%	0%	-4%	-8%
Parcel Based, 1 Connector, Signals Permitted	88%	0%	-3%	-8%
Parcel Based, 2 Connectors, Signals not Permitted	66%	-7%	-8%	-11%
Parcel Based, 2 Connectors, Signals not Permitted "Micromanaged" Demand	87%	-5%	-9%	-12%
Parcel Based, 2 Connectors, Signals Permitted	54%	-6%	-8%	-10%
Parcel Based, 4 Connectors, Signals not Permitted	32%	-9%	-9%	-9%
Parcel Based, 4 Connectors, Signals Permitted	23%	-9%	-10%	-9%

As shown in Table 5.2, all parcel-based methodologies increased the resultant flow on local streets when compared against the original centroid connector structure. In addition, a marginal decrease in the flow on the higher capacity links in the network is evident. The largest increases in flow on local streets were in the one connector cases and the two connector case where the demand is assigned to a specific connector as a function of parcel density ("micromanaged" demand). Not coincidentally, these were the three cases where the selection of a connector by a simulated vehicle was not a result of shortest path assignment; thus, the vehicles were likely forced to enter the network closer to the zonal centroid than the shortest path algorithm would have optimally selected. Aside from the significant increase in flow utilizing the lowest capacity roadways with application of the parcel-based methodology, which decreased as the number of connectors increased, there wasn't a lot of variation amongst the scenarios for collector streets, minor arterials, or principal arterials; in other words, the changes in flow for the collector streets, minor arterials, and principle arterials were fairly stable with respect to the parcel-based methodology for centroid connector placement.

Table 5.3: Changes in V/C Ratio (number of links)

	≥ 1.0	$0.75 \leq V/C < 1.0$	$0.5 \leq V/C < 0.75$	$0.5 \leq V/C < 0$	$= 0$
Base	173	188	257	1438	551
Parcel Based, 1 Connector, Signals not Permitted	202	178	370	1524	334
Parcel Based, 1 Connector, Signals Permitted	197	184	366	1512	349
Parcel Based, 2 Connectors, Signals not Permitted	176	191	321	1621	299
Parcel Based, 2 Connectors, Signals not Permitted "Micromanaged" Demand	187	201	344	1656	220
Parcel Based, 2 Connectors, Signals Permitted	178	192	317	1594	327
Parcel Based, 4 Connectors, Signals not Permitted	176	178	290	1630	334
Parcel Based, 4 Connectors, Signals Permitted	176	178	278	1627	349

As indicated in Tables 5.3 and 5.4, the number of links with V/C ratio greater than 1 increased while the number of links with V/C ratios equal to 0 decreased. Although V/C >1 is an outcome in static assignment that researchers desire to limit, as it can indicate artificial congestion is occurring (Qian and Zhang, 2012), when combining insights from Tables 5.2 and 5.3, it's more likely that the lower capacity roads are the links that have increased V/C ratios in the parcel-based methodology model results. For example, although the parcel-based methodology with one connector per subzone, not permitted to be located at signalized intersections, saw an increase in ${}^{V}/_{C}$ by 17 percent compared to the original centroid connector structure, given that the flow on lower capacity links increased by 93 percent and the flow on principal arterials decreased by 8 percent, compared to the base, it's reasonable to conclude that the increased ${}^{V}/_{C}$ did not occur on higher capacity links. The parcel-based methodology did successfully increase the number of links being utilized at user equilibrium by 10 percent or more in all scenarios; this indicates the parcel-based methodology is better distributing the traffic across all links of the network (Table 5.5). The percent change between the base scenario and each of the remaining scenarios was calculated via Equation 17.

Table 5.4: Changes in V/C Ratio (percentage change)

	≥ 1.0	$0.75 \leq V/_C < 1.0$	$0.5 \leq V/_C < 0.75$	$0.5 \leq V/_C < 0$	$= 0$
Base	Base case is the reference case				
Parcel Based, 1 Connector, Signals not Permitted	17%	-5%	44%	6%	-39%
Parcel Based, 1 Connector, Signals Permitted	14%	-2%	42%	5%	-37%
Parcel Based, 2 Connectors, Signals not Permitted	2%	2%	25%	13%	-46%
Parcel Based, 2 Connectors, Signals not Permitted "Micromanaged" Demand	8%	7%	34%	15%	-60%
Parcel Based, 2 Connectors, Signals Permitted	3%	2%	23%	11%	-41%
Parcel Based, 4 Connectors, Signals not Permitted	2%	-5%	13%	13%	-39%
Parcel Based, 4 Connectors, Signals Permitted	2%	-5%	8%	13%	-37%

Table 5.5: Variation in Number of Utilized Links per Scenario

	# used links	% change from base
Base	2,056	(Reference)
Parcel Based 1 Connector Signals not Permitted	2,274	10.6%
Parcel Based 1 Connector Signals not Permitted	2,259	9.9%
Parcel Based 1 Connector Signals Permitted	2,309	12.3%
Parcel Based 2 Connectors Signals not Permitted	2,388	16.1%
Parcel Based 2 Connectors Signals not Permitted "Micromanaged" Demand	2,281	10.9%
Parcel Based 2 Connectors Signals Permitted	2,274	10.6%
Parcel Based 4 Connectors Signals not Permitted	2,259	9.9%

Root mean squared error (RMSE) and mean average error (MAE) are the performance metrics used to evaluate success of the model in producing flows consistent with reality. The root mean squared error is the square root of the mean of the square of all the error. It is calculated by Equation 18:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{18}$$

where n is the sample size, y_i is the ith link field count and $\hat{y}_i$ is the ith link flow predicted by VISTA. The RMSE tends to give a high weight to large errors, and is more useful when large errors are understandable. Thus, the mean average error, which measures the magnitude of the error, is also utilized to characterize the accuracy of the model in producing link flows consistent with field data. The mean average error is calculated by Equation 19:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{19}$$

Table 5.6 details the results of the analysis of the scenarios with respect to field data, with RMSE and MAE calculated by Equations 18 and 19, respectively, and the percent change between the base scenario and each of the respective alternative scenarios calculated by Equation 17. Although the parcel-based scenario seems to be performing better than the original scenarios in the aforementioned analyses, based on the behavioral performance metrics found in the literature (e.g. the results show an increase in the number of links utilized, an increase in the resultant flow on low capacity links, and a decrease in the resultant flow on higher capacity links), Table 5.6 shows that these changes are obsolete with respect to accurately representing real world data as there isn't a sufficient change in the root mean squared error or mean average error of the static traffic assignment counts and the field data across the various scenarios. This indicates that this methodology is not better approximating reality, despite a large variation in flows indicated in Tables 5.2-5.5. In fact, the model performance of the model seems to be slightly derogated, as all parcel-based methodologies, except the scenarios with four connectors per subzone, resulted in some sort of minor increase in the RMSE and MAE. One interesting trend is that as the number of connectors increase, the results better mimic real world data, which is in dissonance with other STA research that states that an increase the number of connectors tends to negatively affect the results because it increases the number of options of connectors that can be utilized in a behaviorally inconsistent manner (Qian and Zhang, 2012). Although the results aren't promising in terms of getting more accurate results from static traffic assignment, this parcel-based methodology does represent a data-driven and transparent approach to centroid connector placement in static traffic assignment that performs similarly to the more ambiguous traditional connector placement methodology.

Table 5.6 Count Data Analysis for STA

	RMSE		MAE	
	RMSE	**% Change**	**MAE**	**% Change**
Base	2154.12	(Reference)	1468.71	(Reference)
Parcel Based 1 Connector Signals not Permitted	2,180	2.41%	1491.52	1.55%
Parcel Based 1 Connector Signals Permitted	2,206	1.18%	1524.42	3.79%
Parcel Based 2 Connectors Signals not Permitted	2,180	1.18%	1491.62	1.56%
Parcel Based 2 Connectors Signals not Permitted "Micromanaged" demand	2,223	3.20%	1538.79	4.77%
Parcel Based 2 Connectors Signals Permitted	2,174	0.92%	1486.49	1.21%
Parcel Based 4 Connectors Signals not Permitted	2,160	0.28%	1467.53	-0.08%
Parcel Based 4 Connectors Signals Permitted	2,149	-0.22%	1455.63	-0.89%

5.3 Dynamic Traffic Assignment Results

This section details the impact of placing connectors in locations consistent with the built environment of the subnetwork captured by parcel data. Section 5.3.1 analyzes the ability of the methodology to capture the built environment, in terms of built square footage assigned to connector nodes, better than traditional centroid connector placement techniques. Section 5.3.2 investigates the ability of the methodology to produce resultant flows that are more consistent with real world count and corridor travel time data.

5.3.1 Demand Allocation and Parcel Density at Connector Points

Given DTA's increased reliance on accurate network representation, this section explores the ability of the methodology to visually capture locations that are likely to be high demand locations without manual refinement. Table 5.7 details the number of centroids and connectors created under each of the proposed scenarios. The number of regular links and nodes remained constant across scenarios. When the demand is split between an inner and outer subzone, additional sub-centroids are required to redistribute the demand accordingly. This number is not exactly double because subnetwork boundary centroids are not split as they represent flow across these points extracted from the regional model and are not associated with a TAZ.

Table 5.7: Modeled Scenarios and their Centroid Connector Structure

Scenario	No. of Centroids	No. of Connectors
Base	399	823
Bilevel, 50/50 Demand Split, Signals not Permitted	715	1,545
Bilevel, 90/10 Demand Split, Signals not Permitted	715	1,545
Parcel Based, 2 Connectors, Signals Permitted	715	1,117
Parcel Based, 2 Connectors, Signals not Permitted	715	1,075
Parcel Based, 4 Connectors, Signals Permitted	715	1,585
Parcel Based, 4 Connectors, Signals not Permitted	715	1,443
Parcel Based, 4 Connectors, Signals not Permitted	715	1,443

One of the novel features of this research is that the demand is split uniquely according to parcel density for each individual TAZ, building on the methodology created by Jafari et al. (2015), who first discovered that allocating demand throughout a TAZ, instead of selecting the n nearest nodes to the zonal centroid (Qian and Zhang, 2012), achieves better network loading patterns. Figure 5.1 shows the distribution of inner demand ratios, as explained in Section 3.4.3, for all 158 TAZs in the network. This seems to indicate, at least for this specific test subnetwork, that it is indeed more realistic to place the majority of demand at entry points closer to the zonal centroid. However, the relatively widespread distribution of ratios also suggests that it is valuable to use an approach that is capable of endogenously selecting an appropriate split.

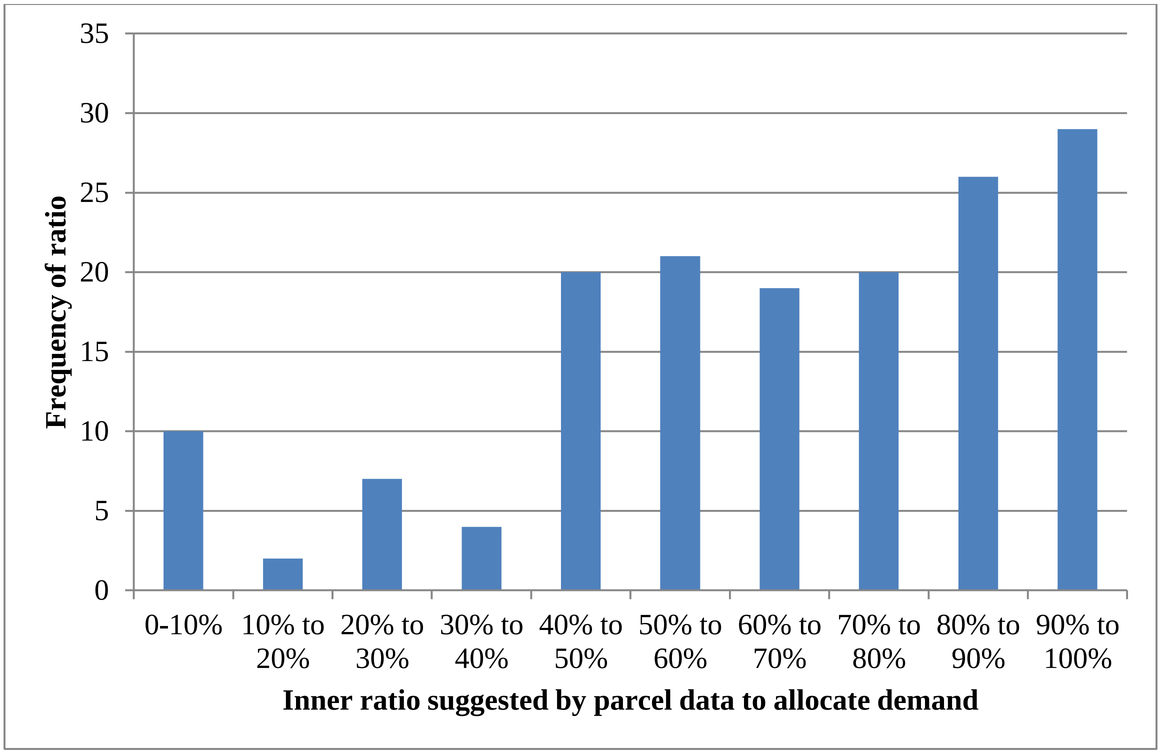


Figure 5.1: Inner Demand Ratio Frequency Based on Parcel Data

As shown in Figure 5.2, the parcel methodology visually appears to be selecting entry/exit nodes that are more consistent with field entry/exit points in the network, with the limitation that only so many connection points can be created with the abstracted model. In order to quantify this feature of the methodology for the entirety of the network, the average parcel density per node and the number of nodes with zero assigned parcel density was calculated and is available in Table 5.8 and Table 5.9, respectively; the percent change between the base scenario and each of the respective alternative scenarios was calculated via Equation 17.

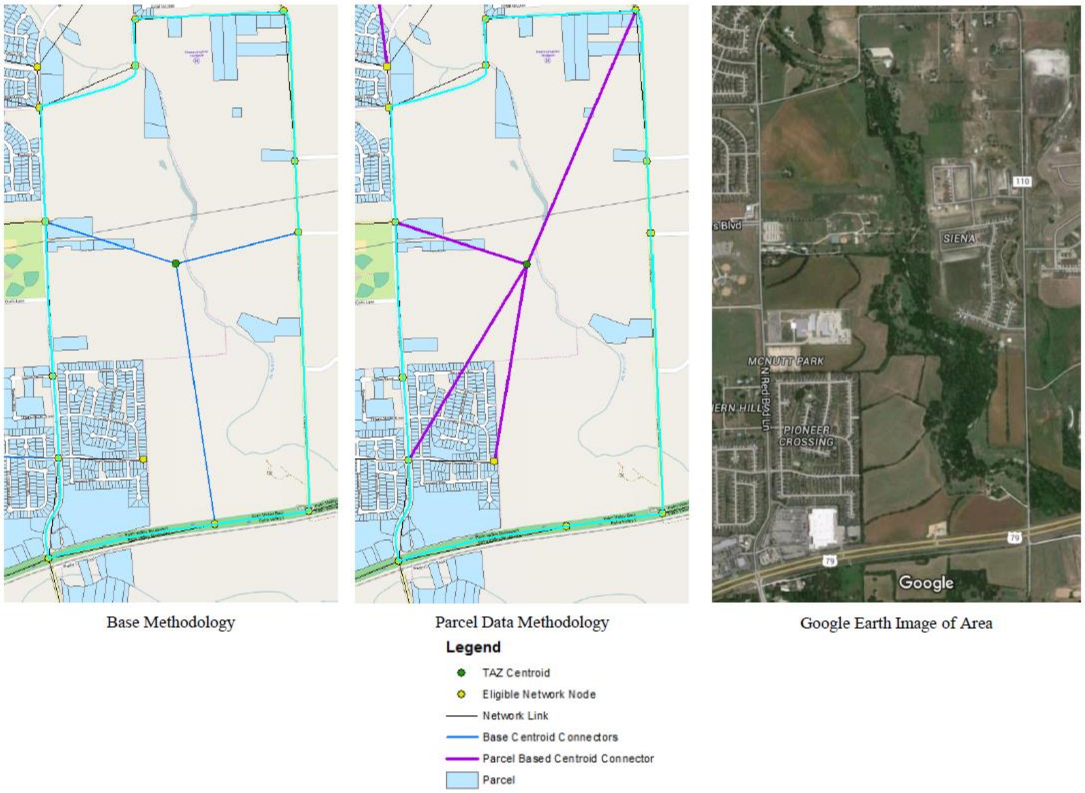


Figure 5.2: Visual Inspection of Centroid Connector Structure

Table 5.8: Parcel Density per Connector Compared Across Scenarios

	Sum of built square footage assigned to entry/exit points	Average built square footage per entry/exit point	% change with respect to base scenario
Base	210,000,000	255,600	(Reference)
Parcel Based 2 Connectors Signals not Permitted	319,000,000	296,500	16.0%
Parcel Based 2 Connectors Signals Permitted	331,000,000	296,700	16.1%
Parcel Based 4 Connectors Signals not Permitted	385,000,000	267,100	4.5%
Parcel Based 4 Connectors Signals Permitted	416,000,000	267,400	4.6%

Table 5.9: Entry/Exit Nodes with Zero Parcels Assigned

	Entry/exit nodes with 0 parcels assigned	% Entry/exit nodes with 0 parcels assigned (as a function of total entry/exit points in the network)
Base	96	11.66%
Parcel Based 2 Connectors Signals not Permitted	76	6.69%
Parcel Based 2 Connectors Signals Permitted	73	6.26%
Parcel Based 4 Connectors Signals not Permitted	104	6.51%
Parcel Based 4 Connectors Signals Permitted	106	5.93%

As can be seen in Table 5.8, the total built square footage assigned to network entry/exit points increases as a function of the utilization of the parcel-based methodology and the number of connectors. In an attempt to prevent a scenario from looking artificially well designed because of an increase in the number of connectors, the average parcel density per entry/exit location was calculated. There's a pretty substantial increase between the base scenario and all four parcel-based approaches in terms of average square footage assigned to each entry/exit location in the network according to Table 5.8, which is promising if built environment is an accurate proxy for demand, as indicated by the literature.

There's also a considerable decrease in the average parcel density at entry/exit locations between the parcel data-driven approach with two connectors and the methodology with four connectors, despite an increase in total square footage assigned to entry/exit points. For example, in the no signalized intersections at connection points scenarios, the average parcel density per entry/exit points drops from 296,500 square feet, with two connectors per subzone, to 267,100 square feet, with four connectors per subzone. This shows that nodes in less developed areas are being selected to meet the "four connector" user input requirement, supporting the need to find a data-informed methodology for endogenously determining optimal connector count for each TAZ or subzone. Somewhat surprisingly, there's no observed large benefit in allowing connectors to be placed at signals, indicating that removing signalized intersections from the list of eligible nodes, as recommended in the literature (Chiu et al., 2011) does not hinder this methodology's ability to capture built environment in the tested network.

Additionally, Table 5.9 explores the frequency of entry/exit nodes being selected at locations where zero parcel data is assigned. For example, 11.66 percent of all of the entry/exit locations in the base network have no built square footage assigned to them;

this is reduced to 5.93 percent of all entry/exit locations in the parcel-basedscenario with four connectors per subzone and signalized intersections permitted as access points. As discussed in Section 3.4.4, instances of zero square footage assigned to an entry/exit node occur in the new methodology when no parcel data is available for the entire TAZ and the n nearest nodes to the zonal centroid is selected as the entry/exit point. Thus, the parcel data-driven approach offers improvements here as well, with a large decrease in the occurrence of an entry/exit point to the network being created with zero built square footage assigned to the entry node (see Figure 5.2). Consistent with other results, there's very little change in the number of entry/exit points with zero assigned built square footage for two and four connector scenarios and only a marginal improvement between signals and no signals. This, again, supports that the removal of signalized nodes from the list of entry/exit points to the network does not have a detrimental impact on the methodology. Additionally, the aforementioned results indicate that simply increasing the number of connectors per TAZ does not have significant impact on the number of connection points with zero square feet assigned for a subnetwork, once again supporting the need for a data-informed approach to select the number of connectors per TAZ or subzone.

5.3.2 Model Performance

Table 5.10 reports a summary of the average system level performance metrics. The lowest total system travel time, total vehicle miles traveled, average origin-destination (OD) travel time, average path link, and the highest average speed are achieved using the parcel data methodology with four connectors per subzone. This strongly suggests that this approach avoids artificial bottleneck creation. This is substantiated through the evaluation of model performance relative to field travel times.

Table 5.10: Test Network Statistics after Convergence

	TSTT (h)	VMT (veh mi)	Average OD Travel Time (min)	Average Speed (mph)	Average Path Length (mi)
Base	22,858	830,970	10.11	36.80	6.24
Bilevel 50/50 Demand Split Signals not Permitted	23,853	847,555	10.56	36.43	6.32
Bilevel 90/10 Demand Split Signals not Permitted	23,911	848,571	10.61	36.25	6.34
Parcel Based 2 Connectors Signals Permitted	21,267	844,503	9.41	37.96	6.30
Parcel Based 2 Connectors Signals not Permitted	21,643	847,728	9.58	37.63	6.33
Parcel Based 4 Connectors Signals Permitted	20,463	820,166	9.05	38.44	6.14
Parcel Based 4 Connectors Signals not Permitted	20,703	826,481	9.16	38.21	6.18

Table 5.11 shows minimum, maximum, and average corridor travel time error for each of the seven strategies. Field travel times were collected along select corridors during peak periods, while model travel times along the same corridors are computed based on the travel time of simulated probe vehicles. The travel time error is the absolute value of the deviation of the model corridor travel time from the field collected corridor travel time. The minimum travel time error is the smallest deviation from the field data that occurred on the 18 corridors where data were available; likewise, the maximum travel time error is the absolute value of the largest deviation between the field data and the modeled data for the 18 corridors. The average travel time error is the average of all

the deviations of modeled travel time on each of the 18 corridors from the field data. The Table 5.11 also shows the percent change between the base methodology for allocating centroid connectors and the remaining six strategies. The percentage change between each respective alternative scenario and the base scenario, with a centroid connector structure created by CAMPO, is calculated in accordance to Equation 17.

Table 5.11: Corridor Travel Time Validation Results

Network	Min TT Error (min)	Max TT Error (min)	Average TT Error (min)	% Change Min	% Change Max	% Change Average
Base	0.028	9.885	1.706	(Reference)	(Reference)	(Reference)
Bilevel, 50/50 Demand Split, Signals not Permitted	0.029	8.094	1.636	5%	-18%	-4%
Bilevel, 90/10 Demand Split, Signals not Permitted	0.029	10.933	1.656	5%	11%	-3%
Parcel Based, 2 Connectors, Signals Permitted	0.039	2.232	0.969	40%	-77%	-43%
Parcel Based, 2 Connectors, Signals not Permitted	0.001	2.173	0.960	-95%	-78%	-44%
Parcel Based, 4 Connectors, Signals Permitted	0.010	2.235	0.982	-65%	-77%	-42%
Parcel Based, 4 Connectors, Signals not Permitted	0.004	2.213	0.973	-85%	-78%	-43%

As one can see, it is clear that the parcel-based performs better than the base methodology with respect to travel time data on the test network. The approach that best matched the field data was the strategy that involved splitting demand and placing connectors via parcel data with signalized intersections not on the eligible node list and two connectors selected per subzone. This resulted in a 95 percent decrease from the base case in minimum corridor travel time error, a 78 percent decrease from the base case in maximum corridor travel time error, and a 44 percent decrease from the base case in average corridor travel time errors. One important observation is that the scenario that performed the best with respect to field corridor travel times is the scenario that achieved the highest average parcel density per entry/exit network node, discussed in Table 5.8, indicating that parcel density is a great proxy for demand generation in the real world.

Table 5.12 shows the results of the error associated with link volume counts. The table also details the percent change between the calibrated base methodology and the remaining six strategies in terms of root mean squared error and mean absolute error, as calculated by Equations 18 and 19, respectively. Field counts, available at various levels of temporal aggregation, were compared with model volumes aggregated in a consistent manner. The percent change for each alternative scenario with respect to the base scenario is calculated via Equation 17.

Table 5.12: Field Traffic Count Validation Results

Network	RMSE	MAE	% Change from Base	
Base	413.9	256.5	---	---
Bilevel, 50/50 Demand Split, Signals not Permitted	440.6	258.5	6%	1%
Bilevel, 90/10 Demand Split, Signals not Permitted	424.3	253.0	2%	-1%
Parcel Based, 2 Connectors, Signals Permitted	367.4	237.3	-11%	-7%
Parcel Based, 2 Connectors, Signals not Permitted	362.6	232.8	-12%	-9%
Parcel Based, 4 Connectors, Signals Permitted	372.2	239.5	-10%	-7%
Parcel Based, 4 Connectors, Signals not Permitted	367.7	234.6	-11%	-9%

Both the root mean square error and the mean average error are consistently lower using the parcel-data based approach when compared against the base model. It is interesting to note that the bi-level approach actually increased the RMSE of the link count errors, though it demonstrated improvement in terms of travel times and resulted in more behaviorally consistent utilization of local links (Jafari et al, 2015). This is likely a consequence of the "one size fits all" approach to distributing demand between the inner and outer subzone. It is clearly shown in Figure 5.1 that while the majority of TAZs support larger inner-to-outer demand split ratios, there is variability across the network.

Much of the data used in this experiment were collected on major streets, and thus, comparatively accurate model results depend on proper allocation of demand along connecting roadways. Parcel level data seems to support a more realistic placement of centroid connectors, the subsequent distribution, and loading of demand for use in DTA. This ultimately reduces the observed errors associated with resultant model flow and simulated travel times on key corridors.

Chapter 6: Discussion

The results presented in the previous section suggest that parcel data may be utilized to produce more accurate static and dynamic traffic assignment model results. Additional insights from this research that may inform future model improvements and implementation are discussed below.

Simply showing that results are more behaviorally consistent in static traffic assignment and dynamic traffic assignment is not a sufficient criterion for establishing that a centroid connector placement strategy is superior to another methodology. As discussed in Table 5.2, the utilization of built environment data to inform the placement of centroid connectors in static traffic assignment significantly increased the utilization of lower capacity links while marginally decreasing the resultant flows on higher capacity links. According to the literature concerning centroid connector placement, this indicates that this new methodology provides a more accurate way to place centroid connectors in static traffic assignment. However, when comparing modeled output with field counts, it is evident that the new methodology did not make a significant contribution to STA's ability to produce results that match count data. Given the inner-to-outer demand split ratios (Figure 5.1), it is hypothesized that the new methodology forced a larger quantity of vehicles to enter the network closer to the centroid of the TAZ than in the base scenario. In the scenarios where the number of connectors is lower, or the demand is forced to enter at a certain location, it is anticipated that this created artificial congestion and encouraged the vehicles selecting a route based on the shortest path algorithm to spread out and utilize an increased number of lower capacity links in their trip closest to the centroid. Ultimately, this methodology cannot mitigate the fundamentally unrealistic low travel time estimations calculated by using link performance functions during congested conditions (Section 2.5.1); thus, the vehicles continued to overload the higher capacity links with the fastest travel times to make the majority of their trip, which is inconsistent with field data.

Areas where most intersections are signalized require special attention. Signalized intersections were found to present a special challenge in the selection of nodes to load demand. In the modeled network there are 150 signalized intersections. When centroid connectors were linked via highest weighted node not excluding these intersections, 50 and 70 signalized intersections were selected as entry nodes for the two and four connector scenarios, respectively. However, when the average built square footage per connector node was examined, no negative implications were observed. One potential mitigation technique is that additional nodes could be created near signalized intersections as possible entry/exit nodes to address this issue; however, the simulation technique used in the selected model application is prone to generating artificial congestion along short links, particularly at signal approaches. Thus, further research is required to address this issue appropriately.

More connectors do not necessarily mean better results. As noted in Table 5.10, though the four connector per subzone scenario seemed to yield the best system performance metrics—total system travel time, vehicle miles traveled, average OD travel time, average speed, and average length path—it was not the scenario that best matched field data. The two-connector-per-subzone scenario marginally outperformed the four-connector scenarios. This supports similar conclusions that more connectors per TAZ does not necessarily mean better results, found to be true in static traffic assignment (Qian and Zhang, 2012). However, as a whole, the results for the two- and four-connector scenarios were very similar and both showed improvements over traditional approaches in a dynamic setting.

Splitting demand arbitrarily within TAZs may not provide accurate loading patterns. The work performed by Jafari et al. (2015) was an important step towards recognizing the importance of centroid connector placement on ensuring reasonable traffic patterns, and providing scalable solutions to improve large regional network models. However, the implemented one-size-fits-all demand split, along with the lack of information regarding the actual location of activities within a TAZ, resulted in limitations in the effectiveness of the approach. Given that this new methodology performs better than the bi-level approaches in terms of capturing realistic travel times and count data in model outputs, it's apparent that the built environment detail provided by the parcel data is paramount to the success of this effort.

Utilizing parcel data to assign network entry/exit locations better approximates real world entry/exit points. Figure 6.1 depicts the centroid connector structure resulting from both a typical placement strategy and the parcel-based approach, along with the location of built environment based on parcel data within a TAZ in the Williamson County network. Figure 6.2 is a Google Earth aerial image of the area depicted in Figure 6.1 to show what the TAZ physically looks like compared to its abstraction by parcel data and the network representation. With the base methodology, two of the three connectors are placed at nodes in the network where there is likely no demand generated (no development indicated in the parcel map or the Google Earth image). Additionally, there is no connector servicing a residential development in the lower left quadrant of the TAZ. Thus, the use of parcel data helps to more accurately load/unload demand at appropriate locations.

The need for a data-driven approach to endogenously select the optimal number of connectors per TAZ is evident. The lack of improvement between the two- and four-connector scenarios, with respect to the average parcel density per connection point and the number of connection points with no built square footage assigned, indicates that the algorithm was forced to select non-ideal nodes as a result of the user specified number of connectors per TAZ. This insight supports the need for future research into how to decide the number of connectors per TAZ and is consistent with the results analyzed with respect to travel time and counts.

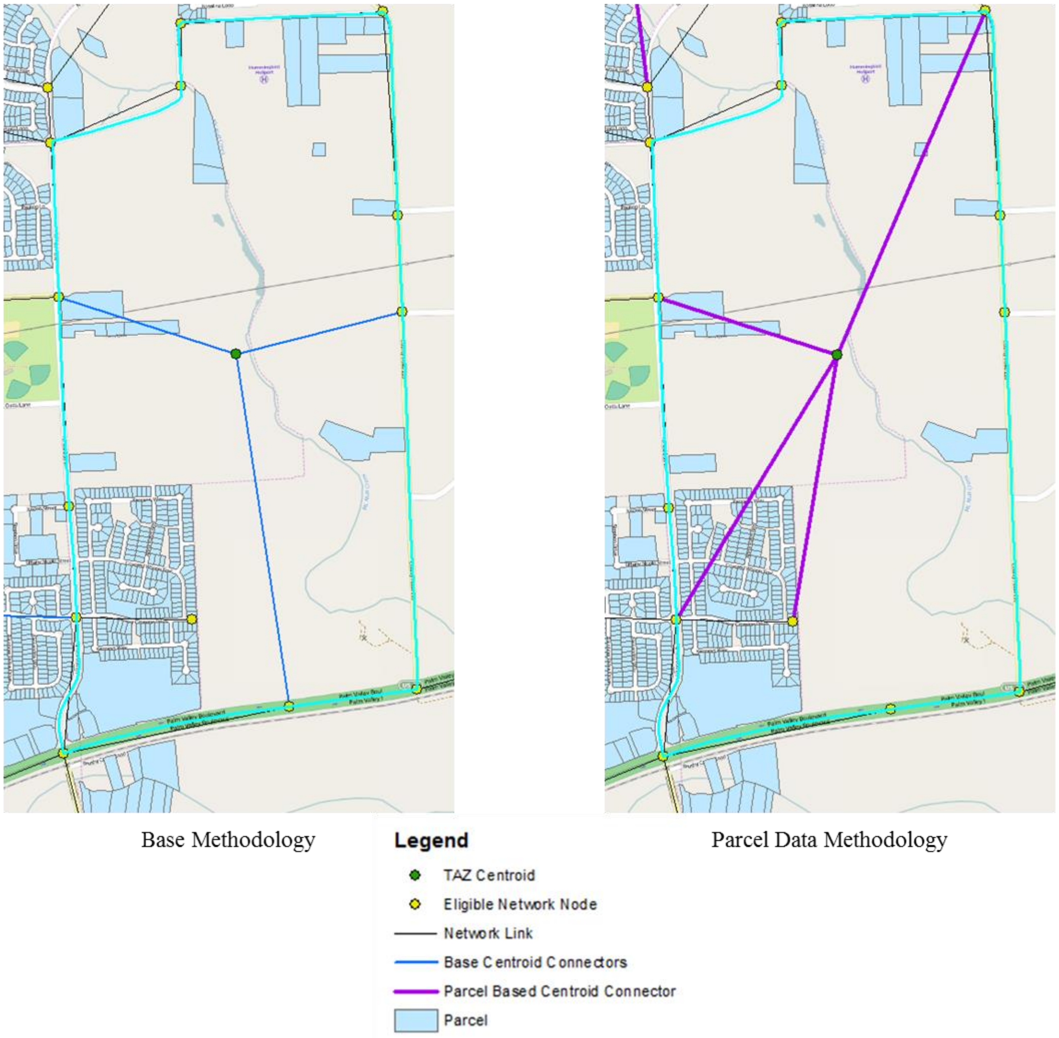


Figure 6.1: Centroid Connector Placement Before and After Comparison

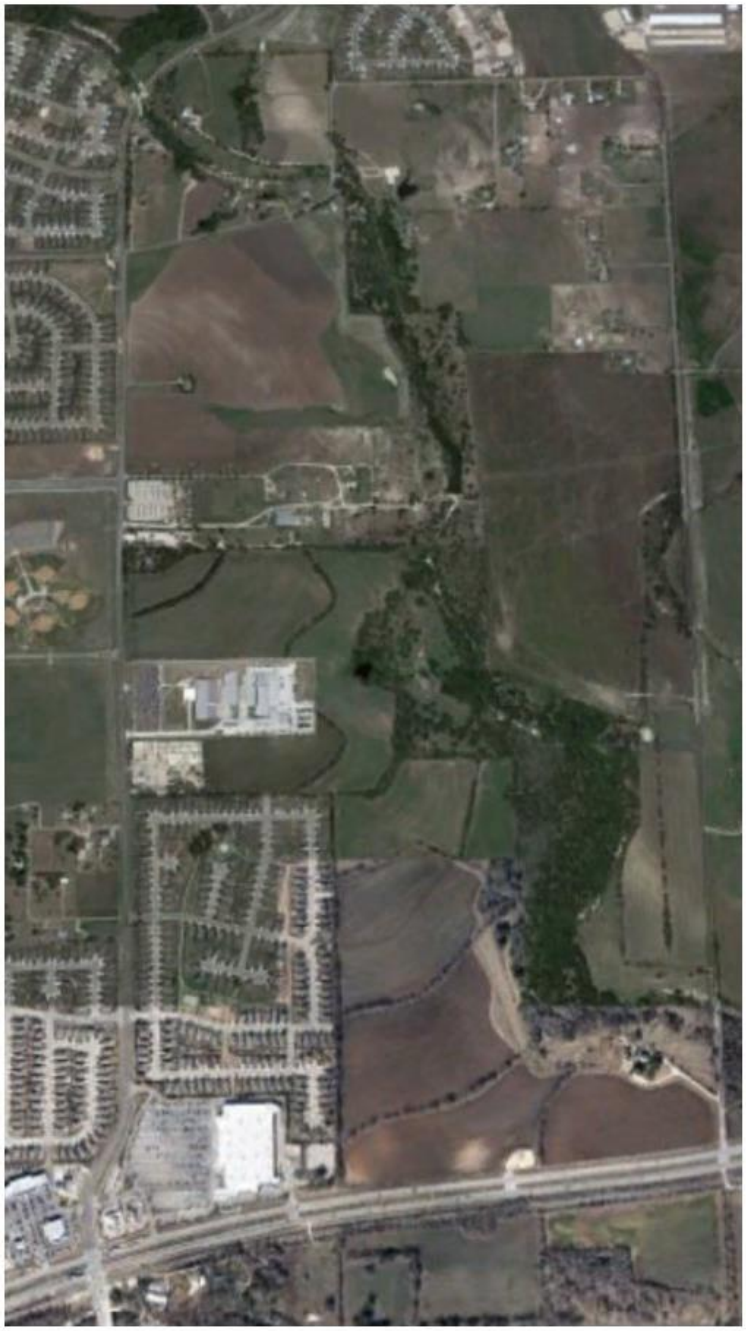


Figure 6.2: Google Earth Image of Area Modeled in Network in Figure 6.1

Chapter 7: Conclusions and Recommendations

This thesis proposes a data-driven methodology informed by parcel level built environment data to automatically place centroid connectors in networks for both static and dynamic traffic assignment applications. Simulation-based DTA models are particularly sensitive to the topological detail of the traffic network, including the location of centroid connectors. Traditional centroid connector placement strategies may lead to excessive congestion and unrealistic traffic patterns, while manual network refinement is prohibitive in large regional models. This research uses parcel-level data to both allocate travel demand between two sub-areas within each TAZ and to select appropriate network access points within each of these zones. It extends previous work by allowing the demand split among TAZ sub-areas to vary across zones, and by considering the parcel density when selecting network entry/exit locations.

7.1 Implications of Results

Static Traffic Assignment

Although the results indicated that the data-driven methodology for centroid connector placement by parcel data achieve more behaviorally consistent resultant flows, no significant improvements were observed with respect to matching field count data. It is hypothesized that the strategy encourages vehicles to make use of a larger variety of lower functional class links by forcing vehicles to enter the network at a limited number of locations closer to the zonal centroid. However, this ultimately did nothing to discourage vehicles from using higher functional class links to make the majority of their trip, as link performance functions make the highest functional class links unrealistically appealing under the principle of user equilibrium. However, this thesis presents a simple, transparent, and data-driven approach for centroid connector placement in static traffic assignment that performs as well as traditional methods. The systematic approach presented in this thesis, provides a more robust and scientific approach to centroid connector placement and further research can potentially yield improved static modeling.

Dynamic Traffic Assignment

Numerical experiments suggest that the proposed methodology leads to solutions that are more consistent with field data than both traditional centroid placement approaches and previous research findings. In the numerical experiments conducted on a real-world network, the approach involving two connectors per subzone while avoiding signalized intersections produced the most realistic results. When compared against real world travel times on 18 corridors in the network, the maximum travel time error was reduced to just 2 minutes and 11 seconds and the average travel time error dropped to under a minute (58 s). This was a 95 percent and a 44 percent improvement, respectively, when compared against the base network. Link counts were also found to be more consistent with real-world data when the data-driven approach to centroid connector placement and demand split was used. The RMSE of the 1,305 links with field traffic counts was found to decrease by 12 percent compared to the base case.

The findings also suggest that a larger number of centroid connectors does not necessarily lead to better model results, verifying what had been suggested in the literature. The presence of traffic signals at intersections was observed to introduce additional challenges in the placement of connectors, which may motivate further research, though it was shown that avoiding the placement of connectors at these locations improved the results. In summary, the results are encouraging and highlight the value and importance of collecting, processing and understanding new data sources in the development of traffic models.

7.2 FUTURE RESEARCH

This thesis lays the groundwork for future research extensions and considerations for centroid connector placement in traffic assignment. The next steps in this research effort are to refine how demand is allocated across connectors and explore methods to endogenously determine an appropriate number of centroid connectors to generate per zone. Using parcel density to determine the appropriate number of connectors per TAZ, eliminating another user defined input, is anticipated to not only supplement this process, but to further automate the implemented procedure.

References

Balakrishna, R., Ben-Akiva, M., Bottom, J., & Gao, S. (2013). Information impacts on traveler behavior and network performance: State of knowledge and future directions. In V. S. Ukkusuri & K. Ozbay (Eds.), *Advances in Dynamic Network Modeling in Complex Transportation Systems* (pp. 193–224). New York, NY: Springer New York. Retrieved from http://dx.doi.org/10.1007/978-1-4614-6243-9_8

Beckmann, M. J., McGuire, C. B., & Winston, C. B. (1956). *Studies in the economics of transportation*. Connecticut: Yale University Press.

Ben-Akiva, M. E., Koutsopoulos, H. N., Mishalani, R. G., & Yang, Q. (1997). Simulation laboratory for evaluating dynamic traffic management systems. *Journal of Transportation Engineering*, *123*(4), 283–289. http://doi.org/10.1061/(ASCE)0733-947X(1997)123:4(283)

Benezech, V., & Leurent, F. (2013). Equilibrium traffic assignment: A new model for spatially disaggregate demand. Presented at the Annual Meeting of the Transportation Research Board.

Boyce, D., Lee, D.-H., & Ran, B. (2001). Analytical models of the dynamic traffic assignment problem. *Networks and Spatial Economics*, *1*(3–4), 377–390. http://doi.org/10.1023/A:1012852413469

Boyce, D. E., Ran, B., & Lebanc, L. J. (1995). Solving an instantaneous dynamic user-optimal route choice model. *Transportation Science*, *29*(2), 128–142. http://doi.org/10.1287/trsc.29.2.128

Boyles, S., Ukkusuri, S. V., Waller, S. T., & Kockelman. (2006). Comparison of static and dynamic traffic assignment under tolls: A study of the Dallas-Fort Worth network. Presented at the 85th Annual Meeting of the Transportation Research Board, Washington, D. C. Retrieved from http://www.ce.utexas.edu/prof/kockelman/public_html/trb06transcaddta.pdf

Buckley, S., & Lightman, D. (2015). Ready or not, big data is coming to a city (transportation agency) near you. Presented at the 94th Annual Meeting of the Transportation Research Board, Washington, D. C.

Burt, M., Cuddy, M., & Razo, M. (2014). *Big data's implication for transportation operations: An exploration* (No. Publication No. FHWA-JPO-14-157). Cambridge, MA: John A. Volpe National Transportation Systems Center.

Cambridge Systemmatics. (2015). *Status of activity-based models and dynamic traffic assignment at peer MPOs*. Cambridge Systematics, Inc. Retrieved from https://www.mwcog.org/uploads/committee-documents/aVxfXVhW20150827091119.pdf

Carey, M. (1992). Nonconvexity of the dynamic traffic assignment problem. *Transportation Research, 26B*(2), 127–133.

Carey, M., & Subrahmanian, E. (2000). An approach to modelling time-varying flows on congested networks. *Transportation Research Part B: Methodological*, *34*(3), 157–183. http://doi.org/10.1016/S0191-2615(99)00019-3

Chen, H.-K., & Hsueh, C.-F. (1998). A model and an algorithm for the dynamic user-optimal route choice problem. *Transportation Research Part B: Methodological*, *32*(3), 219–234. http://doi.org/10.1016/S0191-2615(97)00026-X

Chiu, Y.-C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, S. T., & Hicks, J. (2011). *Dynamic traffic assignment: A primer* (No. Number E-C153). Retrieved from http://onlinepubs.trb.org/onlinepubs/circulars/ec153.pdf

Corthout, R., Flötteröd, G., Viti, F., & Tampère, C. M. J. (2012). Non-unique flows in macroscopic first-order intersection models. *Transportation Research Part B: Methodological*, *46*(3), 343–359. http://doi.org/10.1016/j.trb.2011.10.011

Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, *28*(4), 269–287. http://doi.org/10.1016/0191-2615(94)90002-7

Daganzo, C. F. (1995). The cell transportation model, part II: Network traffic. *Transportation Research Part B: Methodological*, *29*(2), 79–93.

Dong, H., Wu, M., Ding, X., Chu, L., Jia, L., Qin, Y., & Zhou, X. (2015). Traffic zone division based on big data from mobile phone base stations. *Transportation Research Part C: Emerging Technologies*, *58, Part B*, 278–291. http://doi.org/10.1016/j.trc.2015.06.007

Duthie, J. C., Nezamuddin, N., Ruiz Juri, N., Rambha, T., Melson, C., Pool, C. M., … Shah, R. (2013). *Investigating regional dynamic traffic assignment modeling for improved bottleneck analysis* (No. FHWA/TX-13/0-6657-1) (p. 81). Center for Transportation Research. Retrieved from http://library.ctr.utexas.edu/ctr-publications/0-6657-1.pdf

Friedrich, M., & Galster, M. (2009). Methods for generating connectors in transport planning models. *Transportation Research Record: Journal of the Transportation Research Board*, *2132*, 133–142. http://doi.org/10.3141/2132-15

Friesz, T. L., Bernstein, D., Smith, T. E., Tobin, R. L., & Wie, B. W. (1993). A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research*, *41*(1), 179–191. http://doi.org/10.1287/opre.41.1.179

Friesz, T. L., Luque, J., Tobin, R. L., & Wie, B.-W. (1989). Dynamic network traffic assignment considered as a continuous time optimal control problem. *Operations Research*, *37*(6), 893–901. http://doi.org/10.1287/opre.37.6.893

Griesenbeck, B. (2006). Preparing parcel-level input data for the activity-based travel model in Sacramento. In *Innovations in travel demand modeling: summary of a conference*. Retrieved from http://onlinepubs.trb.org/onlinepubs/conf/CP42.pdf

Jafari, E., Gemar, M. D., Ruiz Juri, N., & Duthie, J. (2015). Investigation of centroid connector placement for advanced traffic assignment models with added network detail. *Transportation Research Record: Journal of the Transportation Research Board*, *2498*, 19–26. http://doi.org/10.3141/2498-03

Jayakrishnan, R., Mahmassani, H. S., & Hu, T.-Y. (1994). An evaluation tool for advanced traffic information and management systems in urban networks. *Transportation Research Part C: Emerging Technologies*, *2*(3), 129–147. http://doi.org/10.1016/0968-090X(94)90005-1

Lighthill, M. J., & Whitham, G. B. (1955). On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London A: Mathematical,* *Physical and Engineering Sciences*, *229*(1178), 317–345. http://doi.org/10.1098/rspa.1955.0089

Mahmassani, H. S., & Peeta, S. (1995). System optimal dynamic assignment for electronic route guidance in a congested traffic network. In *Urban Traffic Networks: Dynamic Flow Modeling and Control* (pp. 2–27). Berlin: Springer-Verlag.

McNally, M. G. (2007). The four step model. In Hensher (Ed.), *Handbook of Transport Modeling* (2nd ed.). Pergamon. Retrieved from http://www.its.uci.edu/its/publications/papers/CASA/UCI-ITS-AS-WP-07-2.pdf

Merchant, D. K., & Nemhauser, G. L. (1978a). A model and an algorithm for the dynamic traffic assignment problems. *Transportation Science*, *12*(3), 183–199.

Merchant, D. K., & Nemhauser, G. L. (1978b). Optimality conditions for a dynamic traffic assignment model. *Transportation Science*, *12*(3), 200–207.

Morales, M. (2010). *Parcel-level methodolog for estmating commercial, industrial, and institutional water use* (Master's thesis). University of Florida. Retrieved from http://www.conservefloridawater.org/publications/m.morales_thesis.pdf

Nagurney, A. (1998). *Network economics: A variational inequality approach*. Boston: Kluwer Academic Publishers.

National Research Council (Committee on Land Parcel Databases). (2007). *National land parcel data: A vision for the future*. The National Academies Press.

Nezamuddin. (2011). *Improving the efficiency of dynamic traffic assignment through computational methods based on combinatorial algorithm* (Doctoral dissertation). The University of Texas at Austin. Retrieved from https://repositories.lib.utexas.edu/bitstream/handle/2152/ETD-UT-2011-08-4124/NEZAMUDDIN-DISSERTATION.pdf?sequence=1

Organization for Economic Co-opertation and Development. (2014). Big data in transport: Applications, implications, and limitations. In *Annual Summit: Transport for a Changing World Session Summaries*. Retrieved from https://issuu.com/00716/docs/2014-summit-session-highlights

Patriksson, M. (2015). *The traffic assignment problem: Models and methods*. Mineola, New York: Dover Publications.

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. http://doi.org/10.1023/A:1012827724856

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. http://doi.org/10.1016/j.trb.2012.07.006

Ran, B., Boyce, D. E., & LeBlanc, L. J. (1993). A New class of instantaneous dynamic user-optimal traffic assignment models. *Operations Research*, *41*(1), 192–202. http://doi.org/10.1287/opre.41.1.192

Ran, B., & Shimazaki, T. (1989). A general model and algorithm for the dynamic traffic assignment problems. In *Proceedings of the Fifth World Conference on Transportation Research*. Yokohoma, Japan.

Richards, P. I. (1956). Shock waves on the highway. *Operations Research*, *4*(1), 42–51.

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. Publication No. FHWA-HOP-13-015). Retrieved from http://ops.fhwa.dot.gov/publications/fhwahop13015/fhwahop13015.pdf

Sheffi, Y. (1985). *Urban transportation networks: Equilibrium with mathematical programming methods*. Englewood Clifs, NJ: Prentice-Hall, Inc. Retrieved from http://sheffi.mit.edu/sites/default/files/sheffi_urban_trans_networks.pdf

Shi, Q., & Abdel-Aty, M. (2015). Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. *Transportation Research Part C: Emerging Technologies*, *58*, *Part B*, 380–394. http://doi.org/10.1016/j.trc.2015.02.022

Tampère, C. M. J., Corthout, R., Cattrysse, D., & Immers, L. H. (2011). A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B: Methodological*, *45*(1), 289–309. http://doi.org/10.1016/j.trb.2010.06.004

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, *58*, *Part B*, 162–177. http://doi.org/10.1016/j.trc.2015.04.022

Transportation Research Circular. (2011). *75 years of the fundamental diagram for traffic flow theory* (No. Number E-C149). Retrieved from http://onlinepubs.trb.org/onlinepubs/circulars/ec149.pdf

Yu, W., Park, S., Kim, D. S., & Ko, S.-S. (2015). Arterial road incident detection based on time-moving average method in bluetooth-based wireless vehicle reidentification System. *Journal of Transportation Engineering*, *141*(3), 4014084. http://doi.org/10.1061/(ASCE)TE.1943-5436.0000748

Wardrop, J. G. (1952). Road paper: Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, *1*(3), 325–362. http://doi.org/10.1680/ipeds.1952.11259

Ziliaskopoulos, A. K. (2000). A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation Science*, *34*(1), 37–49. http://doi.org/10.1287/trsc.34.1.37.12281

Ziliaskopoulos, A. K., & Mahmassani, H. S. (1993). Time-dependent, shortest-path algorithm for real-time intelligent vehicle highway system applications. *Transportation Research Record: Journal of the Transportation Research Board*, *1408*, 94–100.

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*(1–6), 427–444. http://doi.org/10.1016/S0968-090X(00)00027-9